TL
1
S6
vol. 68

SAE TRANSACTIONS

Society of Automotive Engineers, Inc.
485 Lexington Avenue New York 17, New York

H. E. Chesebrough
President

Joseph Gilbert
Secretary and General Manager

G. A. Delaney
Treasurer

E. D. Allen
Editor

Kathleen Napier
Assistant Editor

All articles appearing in SAE Transactions are indexed by Engineering Index, Inc.

Copyright 1960, Society of Automotive Engineers, Inc.

Price: $3 to members, $7 to public and college libraries and U. S. Government Agencies, $10 to nonmembers. Foreign: $3 to members, $8 to public and college libraries, $11 to nonmembers.

Photolithographed reproductions are available of any article published in SAE Transactions. Prices, on quantities of 100 copies or more, varying with number of pages and illustrations, can be obtained upon request from SAE Special Publications Department, 485 Lexington Avenue, New York 17, N. Y.

21 June 72

Contents

Vol. 68

1960

Other SAE Publications 4

The Machining of Ultra-Strength Alloys — John Maranchik, Jr., J. V. Gould, and P. R. Arzt 5

Radioisotopes in Research — W. P. Evans and E. W. Landen 19

Basic Design of Turbochargers for Diesel Engines — Hans Egli 29

Relaxation of Residual Stresses due to Fatigue Loading — JoDean Morrow, A. S. Ross, and G. M. Sinclair 40

Rating Transmissions from Highway Requirements and Vehicle Specifications — F. C. Meldola 49

Compressive Failures in Transmission Gearing — G. E. Huffaker .. 53

A Study of Self-Contained Starting Systems for Turbojet and Turboprop Engines — Henry R. Schmider and John H. Ferguson, Jr. 60

Possibility of "Spare Tire Elimination" by Use of Dual Compartment Tires — Walter J. Lee 74

The Expanding Polymer Horizon — J. H. DuBois 77

Electric Propulsion — Measurements with a Small Thrust Plasma Generator — B. Gorowitz and B. W. Harned 85

The Identification and Characterization of Rumble and Thud — E. S. Starkman and W. E. Sytz 93

Rolls-Royce Diesel Engines — R. H. Whiteside 101

Non-Air-Breathing Auxiliary Powerplants — D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor 111

The Effect of Residual Stresses Induced by Strain-Peening upon Fatigue Strength — R. L. Mattson and J. G. Roberts 130

Torsionally Resilient Drive Lines — P. J. Mazziotti 137

Prediction in New Metal Joining Processes — John J. Chyle 143

High-Energy Accelerators — Keith R. Symon 157

Some Development Problems with Large Cryogenic Propellant Systems — Daniel A. Heald 164

Development of Two New Allis-Chalmers Diesel Engines — Hans L. Wittek 169

New Elastomers Break Old Barriers — Ralph P. Schmuckal 193

Engine Rumble — A Barrier to Higher Compression Ratios? — R. F. Stebar, W. M. Wiese, and R. L. Everett 206

Ways of Improving Take-Off and Landing — F. W. Kolk 217

Properties of Silicone Rubber for High-Temperature Static Seals — H. E. Todd and J. F. Miazga 224

Electric Drive for Off-Highway Vehicles — H. J. McLean and H. Vitt 232

Hydrostatic Transmissions — P. C. Mortenson 243

Hydrostatic Transmissions for Vehicles — E. H. Bowers 248

Hydrodynamic Split Torque Transmissions — Peter Bloch and Raymond C. Schneider 257

Application of the Cumulative Fatigue Damage Theory to Practical Problems — W. F. Hofmeister 274

Current Trends in Storable Propellants — J. Silverman and S. A. Greene 283

Radiotracers Reveal Engine Wear during Detergent Oil Filtration Study — Harry Halliwell 288

The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear — B. A. Robbins, P. L. Pinotti, and D. R. Jones 298

Computer-Based Selection of Balanced-Life Automotive Gears — G. J. Huebner 305

Design Procedure for Vibration Isolation on Nonrigid Supporting Structures — Sheldon Rubin 318

Spark-Plug Misfiring — Mechanism Studies — H. P. Julien and R. F. Neblett 328

Brake Drums — Investment for Safety and Profit — Paul G. Hykes and Clarence A. Herman 336

Economics of Heavy-Duty Brake Design and Maintenance — R. K. Super 346

The Development of Refractory Sheet Metal Structures — Alan V. Levy and Saul E. Bramer 352

Road Performance of a Comprex Supercharged Diesel Truck — Max Berchtold and H. P. Gull 367
Electric Propulsion Systems — Victor P. Kovacik and Daniel P. Ross 380
Engine Blocks and their Components in Aluminum Die Casting — A. F. Bauer 385
Preliminary Design Considerations for the Structure of a Trisonic Transport — M. G. Childers 396
The Design of a Truck for Better Ride and Handling — H. O. Flynn . 408
Seal Testing To Establish Quality Control Specifications Can Reduce "Leakers" — Robert L. Dega and James D. Symons 440
Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids — H. E. Deen and C. M. Stendahl 448
Automatic Transmission Fluid Viscosity at Low Temperature and its Effect on Transmission Performance — Theodore W. Selby 457
High-Temperature Materials and their Heat-Treatment for Antifriction Applications — O. W. McMullan 468
Turbine Fuel Thermal Stability, CFR Coker and Flight Evalutions — J. D. Rogers, Jr. 474
Soviet Agricultural Tractors and Equipment — Wayne H. Worthington 483
IH High-Speed Lightweight Diesels — A. F. Dewsberry, C. P. Bozos, and J. B. Reeves, Jr. 501
Reduction of Air Pollution by Control of Emission from Automotive Crankcases — P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 514
Material and Structural Damping for Vibration Control — B. J. Lazan 537
Dynamic Balancing of Rebuilt Engines and Driveline Components — Ralph T. Buscarello 548
Radiation from Flames in Gas Turbine Combustors — R. M. Schirmer and L. A. McReynolds 554
Physical and Chemical Ignition Delay in an Operating Diesel Engine Using the Hot-Motored Technique — Part II — C. W. Chiang, P. S. Myers, and O. A. Uyehara 562
Refractory Nonmetals for Use in Hypersonic Aircraft — John M. Nowak 571
The First Year of the Jet Age — Reflections — F. W. Kolk 578
How Good is Testing? A Correlation of Customer, Laboratory, and Proving Ground Experience — William A. McConnell 588
Nuclear Rockets for Interplanetary Propulsion — Frank E. Rom and Paul G. Johnson 600
Transistor Switched Ignition Systems — G. E. Spaulding, Jr. 610
Correcting Horsepower Output, A Realistic Method for Diesel Engines — A. K. Blackwood and W. J. McCulla 620
Properties of Asbestos Reinforced Laminates at Elevated Temperatures — Norman E. Wahl 627
Radiation Effects Studies for the Development of Radiation-Resistant Aircraft Subsystems — J. J. Tierney 634
Horsepower Rating for Commercial Vehicle Brakes — John H. King and Alex F. Stamm 646
A Universal Means for Rating Diesel Engines for Deposits and Wear — F. A. Robbins, L. G. Schneider, and G. H. Shea 653
Optimum Joint Design for High-Temperature Honeycomb Panels — Frank J. Filippi and Boris Levenetz 658
GMR Stirling Thermal Engine, Part of the Stirling Engine Story — 1960 Chapter — Gregory Flynn, Jr., Worth H. Percival, and F. Earl Heffner 665
Developing Transaxle Fluid — N. A. Hunstad, R. A. Wilkins, R. E. Osborne, and E. D. Davison 685
Radioactive Cylinders — A Tool for Wear Research — W. C. Arnold, V. T. Stonehocker, W. J. Braun, and D. N. Sunderman 699
Index 705

Other SAE Publications

SP-14 Cooperative Testing of Torsiographs and Calibrators

SP-23 Torquing of Nuts in Aircraft Engines

SP-37 Highway Safety — Review and Forecast — 1947 Beecroft Memorial Lecture by Paul G. Hoffman

SP-71 Highway Safety through Eyes of Motor Vehicle Administrator — 1951 Beecroft Memorial Lecture by Rudolph King

SP-72 Traffic Law Enforcement — 1952 Beecroft Memorial Lecture by Franklin M. Kreml

SP-73 Role of Public Information in Safety Cause — 1953 Beecroft Memorial Lecture by W. Earl Hall

SP-74 Driver Education — Key to Safe Operation of Motor Vehicles — 1954 Beecroft Memorial Lecture by Amos E. Neyhart

SP-75 Traffic Engineering, Its Role in Safe and Efficient Operation of Streets and Highways — 1955 Beecroft Memorial Lecture by D. Grant Mickle

SP-76 Skill and Will To Drive Safely — 1956 Beecroft Memorial Lecture by O. E. Hurt

SP-113 Engine Deposits Symposium — engine deposit data, 1939–1952, 18 papers and discussion

SP-116 Engine Wear Symposium — 20 papers

SP-119 Engineering Know-How in Engine Design, Part 1 (1953) — foundamentals of engine design

SP-122 Engineering Know-How in Engine Design, Part 2 (1954) — theory, test data, and empirical formulas of engine design

SP-127 Symposium on Cam and Tappet Wear Problems (1954)

SP-130 Design, Evaluation, and Selection of Heavy-Duty Rear Axles — first L. Ray Buckendale Lecture by K. W. Gordon

SP-131 Role of Turbine in Future Vehicle Powerplants — second L. Ray Buckendale Lecture by C. G. A. Rosen

SP-132 New Drive Lines for New Engines — third L. Ray Buckendale Lecture by W. P. Mitchell

SP-133 Bearing Applications for Heavy-Duty Axles — fourth L. Ray Buckendale Lecture by R. M. Riblet and C. M. Kitson

SP-134 Design of Planetary Gear Trains — fifth L. Ray Buckendale Lecture by O. K. Kelley

SP-136 The SAE Story — history of SAE 1905–1955

SP-137 Engineering Know-How in Engine Design, Part 3 (1955) — reciprocating engine dynamics, fuel applications, two-stroke engines, opposed piston diesel

SP-139 Symposium on Fuels and Lubricants — detergency activity tes s, lubrication of hydraulic valve lifters, air pollution (1955)

SP-140 Symposium on Fuel Injection — 8 papers on passenger-car application

SP-143 Engineering Know-How in Engine Design, Part 4 (1956) — gear design and manufacture, supercharging, development of small engines, automotive gasoline injection, diesel fuel injection

SP-144 Symposium on Methods of Detecting and Measuring Seams — report by Division 30(A) of SAE Iron and Steel Technical Committee

SP-148 Engineering Know-How in Engine Design, Part 5 (1957) — engine crankshafts and connecting rods, bearings and bearing lubrication, pistons and rings

SP-152 Engineering Know-How in Engine Design, Part 6 (1958) — valve train engineering, two-stroke engines, fuel injection, liquid cooling systems, and aircooled engine design criteria

SP-154 Better Truck Ride for Driver and Cargo — ride as affected by suspension system and configuration of vehicle, recommendations of SAE Committee on Ride, report by R. N. Janeway

SP-155 Better Braking — How? (1958)

SP-157 Engine Noise Symposium — combustion-chamber noises in high compression ratio engines, 11 papers (1958)

SP-158 Multifuel Engine Symposium — 1958 report on ordnance requirements and industry progress on multifuel engines

SP-159 Aluminum Engine Symposium (Part 1) — 5 papers and discussion on manufacturing techniques for automotive aluminum engines (1958)

SP-161 Evaluating Automatic Transmission Fluids (1959)

SP-162 Make Slides Worth While

SP-163 Engineering Know-How in Engine Design, Part 7 (1959) — engine roughness, viscous torsional vibration damper, vibration engine mounting, air cleaners, silencers

SP-165 Safety The Motorist Gets — insuring safety in motor cars: design factors, material selection, testing, quality control, reduction of driver fatigue, durability of structural parts (1959)

SP-166 Engine and Speed Governors (1959)

SP-168 Anti-Friction Bearings Seminar — design, materials, lubrication, service

SP-171 Tire and Wheel Runout and Balance Problems

SP-172 Computer-Based Selection of Balanced-Life Automotive Gears — sixth L. Ray Buckendale Lecture by G. J. Huebner

SP-173 Applications and Limitations of New Developments in Metals Joining Processes, by P. J. Rieppel (1960)

SP-174 Automobile Side-Impact Collisions, by D. M. Severy, J. H. Mathewson, and A. W. Siegel (1960)

SP-175 What's New in Engineering Materials for 1960 — technological developments in steel, plastics, rubber, plating, and aluminum

SP-178 Engineering Know-How in Engine Design, Part 8 (1960) — valve train dynamics, dynamometers, radioisotopes, instrumentation for testing, field testing diesels

SP-179 Engineering Know-How in Engine Design — includes SP-119, SP-122, SP-137, SP-143, SP-148, SP-152, SP-163, and SP-178

SP-180 Automobile Aerodynamics — fundamentals and techniques of testing (1960)

SP-181 Surface Rolling Symposium — techniques and mechanical prestressing (1960)

SP-183 B-58 Reliability Program (1960)

SP-214 The Hunt Is on for New Sources of Propulsion Engery — reprint from August, 1957, SAE *Journal*

SP-216 Looking Ahead in Fuels for Automotive Transportation — 1957 Horning Memorial Lecture by J. M. Campbell

SP-218 Single-Cylinder Engine Fuel Research — 1958 Horning Memorial Lecture by A. W. Pope, Jr.

SP-328 1959 Tractor Production Forum — heat treatment, production planning, engineering-purchasing relationships, quality control, automation

SP-329 1959 L.A. Aeronautic Production Forum — propellant handling, missile support equipment, materials handling, cost control, rocket case manufacturing, operations research, optics, reliability, shear forming, sandwich construction, high-temperature coatings

SP-330 1960 National Production Forum — management development, cold extrusion, quality control, high-energy forming, measurement, materials handling, product evaluation

SP-331 1960 Aerospace Manufacturing Engineering Forum — facilities planning, reliability, manufacturing methods and research

the Machining of Ultra-Strength Alloys

John Maranchik, Jr.,

J. V. Gould, and P. R. Arzt

Metcut Research Associates Inc.

THE AIR FORCE has set up a program to evaluate the machining characteristics of the more commonly used high-strength thermal-resistant materials. This paper describes the test results to date.

Four materials are being tested: SAE 4340 quenched and tempered to 50-55 Rockwell C, SAE designation H11 quenched and tempered to 50-55 Rockwell C, AM-350 solution treated and aged, and A-286 solution treated and aged. Machining tests include: turning milling, drilling, and tapping.*

APPLICATION of ultra-high-strength thermal-resistant alloys for airframe, missile, and engine components has been steadily increasing in the past few years. The speed of many of the present-day aircraft imposes thermal conditions on the airframe and powerplant which necessitates the use of ultra-high-strength thermal-resistant alloys for critical components. Future aircraft designed to operate at even higher speeds with increased aero-dynamic friction heating and higher engine temperatures will require an even greater use of such alloys in airframe, missile, and engine construction.

Many of these alloys are new to the industry and meager machining information is available for use in setting up production machining operations. For the more familiar materials such as the martensitic low alloy steels, hot-work die steels, and martensitic stainless steels, there is a lack of data for machining these materials in the high hardness ranges. Hence, a trial-and-error method must be used in the production operations which involves considerable time and expense. The loss of time due to development of the production machining process can seriously disrupt planned schedules for new airframes and engines. Because schedules must be adhered to, production manufacturing departments are limited

* Paper presented at SAE National Aeronautic Meeting, New York, March 31, 1959.

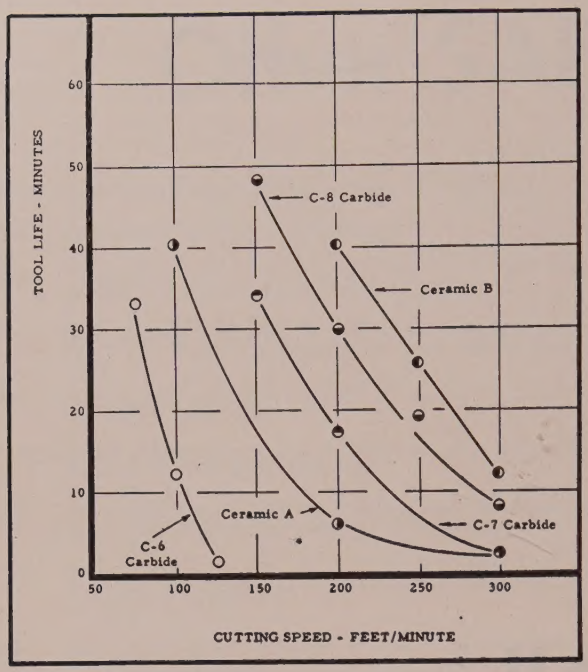


Fig. 1 — Turning 4340 steel quenched and tempered to 514 Bhn (52 Rockwell C). Tool: carbide (see graph); side rake: 5 deg neg.; back rake: 5 deg neg.; nose radius: 0.032 in.; side cutting edge angle: 15 deg; end cutting edge angle: 15 deg; relief: 5 deg. Mechanical chip breaker. Feed: 0.009 in. per rev.; depth of cut: 0.100 in.; cutting fluid: none; wearland: 0.015 in.

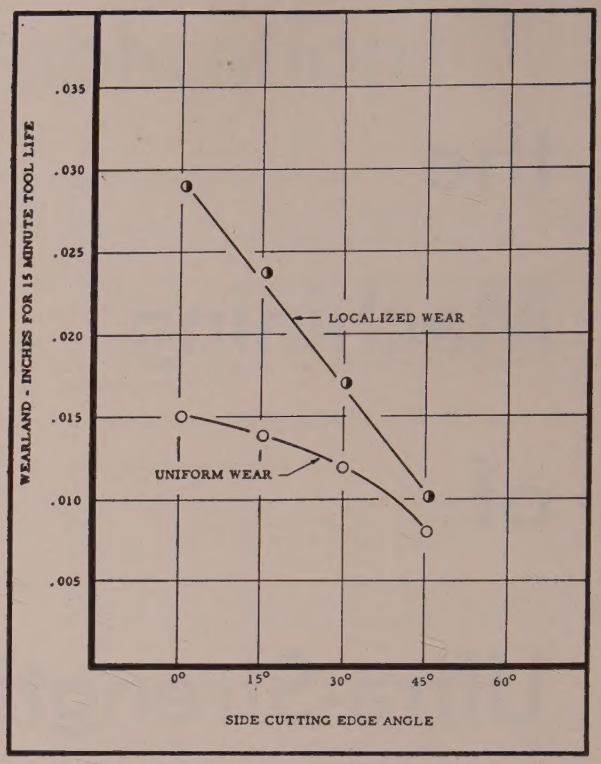


Fig. 3 — Turning 4340 steel quenched and tempered to 514 Bhn (52 Rockwell C); effect of side cutting edge angle on tool wear. Tool: C-7 carbide; side rake: 5 deg neg.; back rake: 5 deg neg.; nose radius: 0.032 in.; side cutting edge angle: 15 deg; end cutting edge angle: 15 deg; relief: 5 deg. Mechanical chip breaker. Feed: 0.009 in. per rev.; cutting speed: 150 fpm; depth 0.100 in.; cutting fluid: none.

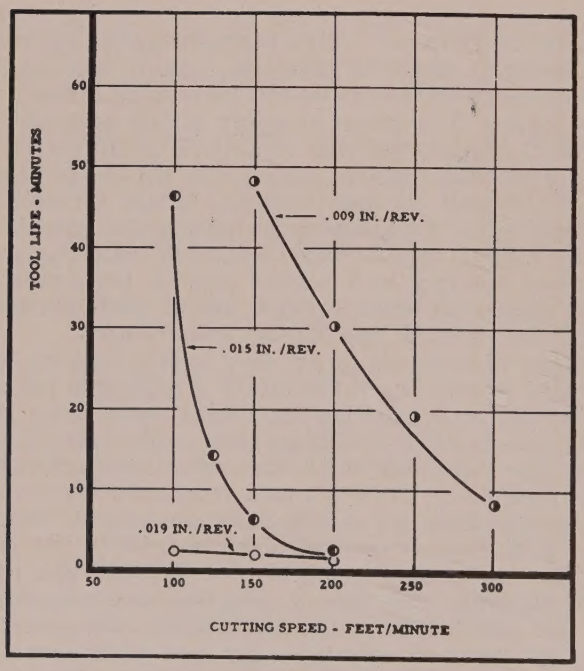


Fig. 2 — Turning 4340 steel quenched and tempered to 514 Bhn (52 Rockwell C); effect of feed. Tool: C-8 carbide; side rake: 5 deg neg.; back rake: 5 deg neg.; nose radius: 0.032 in.; side cutting edge angle: 15 deg; end cutting edge angle: 15 deg; relief: 5 deg. Mechanical chip breaker. Feed: see graph; depth of cut: 0.100 in.; cutting fluid: none; wearland: 0.015 in.

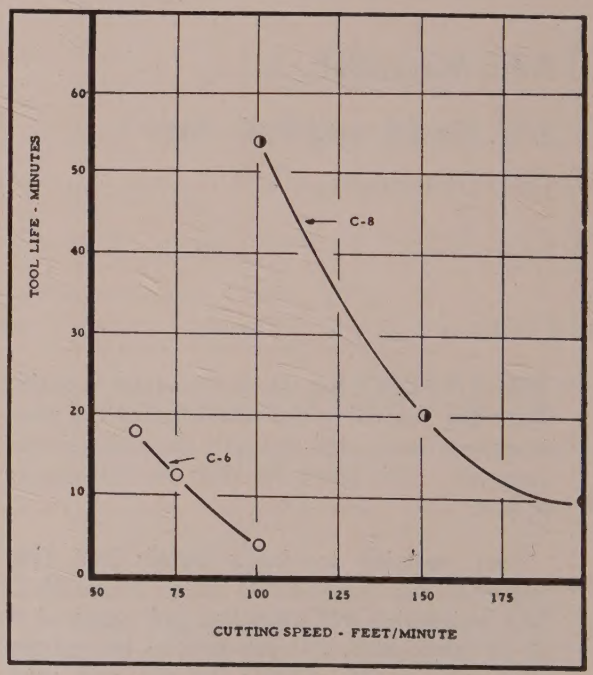


Fig. 4 — Turning H11 quenched and tempered to 514 Bhn (52 Rockwell C); effect of carbide. Tool: carbide (see graph); side rake: 5 deg neg.; back rake: 5 deg neg.; nose radius: 0.032 in.; side cutting edge angle: 15 deg; end cutting edge angle: 15 deg; relief: 5 deg. Mechanical chip breaker. Feed: 0.009 in. per rev.; depth: 0.100 in.; cutting fluid: none; wearland: 0.015 in.

as far as process development is concerned. As a result, manufacturing of the components is done by any available method. The manufacturing process, in many cases, loses the identity of a production operation and becomes a tool room operation. The need for machinability information to be used as an aid in setting up production processing on the high-strength thermal-resistant alloys becomes evident.

In view of the difficulties being encountered in machining these alloys, the Air Force has set up a machinability research program to evaluate the machining characteristics of the more commonly used high-strength thermal-resistant materials. The work is being performed by Metcut Research Associates Inc. under contract to the Wright Aeronautical Division of the Curtiss Wright Corp. under contract No. AF 33 (600)-35967 sponsored by the Air Materiel Command.

Table 1 — Groups of High-Strength Thermal-Resistant Alloys

Alloy Group	Representative Alloy Selected for Machinability Studies
Martensitic low alloy steels	SAE 4340 quenched & tempered to 50–55 R_c
Hot-work die steels	SAE designation H11 quenched & tempered to 50–55 R_c
Precipitation hardening stainless steels	AM-350 solution treated & aged
Austenitic stainless steels	A-286 solution treated & aged

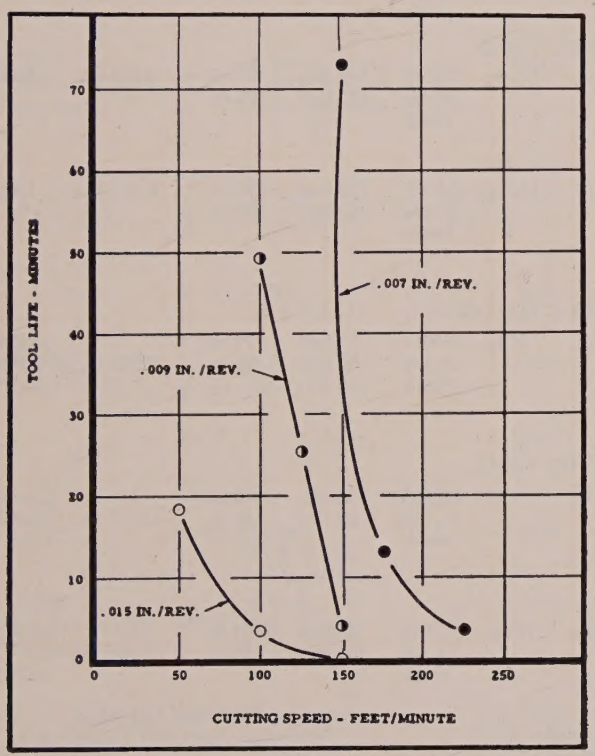


Fig. 5 — Turning H11 quenched and tempered to 514 Bhn (52 Rockwell C); effect of feed rate. Tool: C-8 carbide; side rake: 5 deg neg.; back rake: 5 deg neg.; nose radius: 0.032 in.; side cutting edge angle: 15 deg; end cutting edge angle: 15 deg; relief: 5 deg. Mechanical chip breaker. Feed: see graph; depth: 0.100 in.; cutting fluid: none; wearland: 0.015 in.

This paper will present results of machinability studies made to date on representative high-strength thermal-resistant alloys which can be divided into four groups, as shown in Table 1. The machining information is presented in the form of tool life curves (Figs. 1–32) and summaries of machining data (Table 2, pp. 8–9).

Turning Tests

Turning proved to be the least difficult of the various types of machining operations being studied for the ultra-high-strength thermal-resistant alloys. It was possible to obtain reasonable tool life in turning without resorting to unusual types of tools, tool geometries, or techniques. Good tool life results can be obtained by adhering to the following general recommendations:

1. Use a rigid machine and strong, solid tools and fixtures.
2. Use the proper type of carbide. For a given type of carbide, select the hardest grade that will perform without chipping.
3. Use cutting speeds, feeds, and tool geometries selected from turning data obtained under controlled cutting conditions.

Tool life results for turning the following alloys are presented in Figs. 1–7:

1. SAE 4340 quenched and tempered to 52 R_c

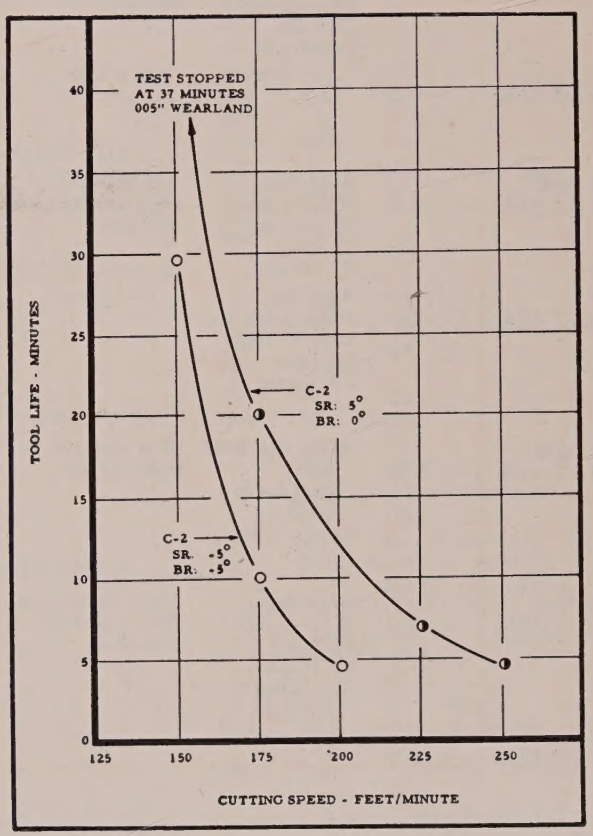


Fig. 6 — Turning AM-350 solution treated and aged to 444 Bhn (47 Rockwell C). Tool: C-2 carbide; side rake: 5 deg (see graph); back rake: 0 deg (see graph); nose radius: 0.032 in.; side cutting edge angle: 15 deg; end cutting edge angle: 15 deg; relief: 5 deg. Mechanical chip breaker. Feed: 0.009 in. per rev.; depth of cut: 0.100 in.; cutting fluid: none; wearland: 0.015 in.

Table 2 — Machining Data for

Material	Tool Material	Tool Geometry	Cutter	Depth of Cut	Width of Cut	Feed	Cutting Speed	Tool Life	Wear-land	Cutting Fluid
TURNING — CARBIDE TOOLS										
SAE 4340 Q & T-52 R_c	C-8	SR: – 5 deg SCEA: 15 deg BR: – 5 deg ECEA: 15 deg Relief: 5 deg	½ in. sq. throwaway holder with mechanical chip breaker	0.100 in.	—	0.009 in. per rev.	150 fpm 300 fpm	47 min 10 min	0.016 in.	None
H11 Q & T-52 R_c	C-8	Same	Same	0.100 in.	—	0.009 in. per rev.	100 fpm 200 fpm	50 min 10 min	0.016 in.	None
AM-350 Solution treated and aged	C-2	SR: 5 deg SCEA: 15 deg BR: 0 deg ECEA: 15 deg Relief: 5 deg	Same	0.100 in.	—	0.009 in. per rev.	150 fpm 200 fpm	60 min 5 min	0.016 in.	None
A-286 Solution treated and aged	C-2	SR: – 5 deg SCEA: 15 deg BR: – 5 deg ECEA: 15 deg Relief: 5 deg	Same	0.100 in.	—	0.009 in. per rev.	150 fpm 250 fpm	25 min 7 min	0.016 in.	None
FACE MILLING — CARBIDE TOOLS										
SAE 4340 Q & T-52 R_c	C-6	AR: 0 deg ECEA: 5 deg RR: – 15 deg CI: 8 deg CA: 45 deg	5 in. diameter, 5-tooth inserted tooth face mill	0.100 in.	2 in.	0.005 in. per tooth	150 fpm 225 fpm	330 in. 180 in.	0.016 in.	None
H11 Q & T-52 R_c	C-2	Same	Same	0.100 in.	2 in.	0.005 in. per tooth	125 fpm 175 fpm	400 in. 100 in.	0.016 in.	None
SIDE MILLING — CARBIDE TOOLS										
SAE 4340 Q & T-52 R_c	C-6	AR: – 5 deg ECEA: 5 deg RR: – 10 deg CI: 8 deg CA: 45 deg	7 in. diameter, 6-tooth inserted tooth face mill	0.100 in.	1¾ in.	0.0075 in. per tooth	150 fpm 220 fpm	380 in. 75 in.	0.016 in.	None
H11 Q & T-52 R_c	C-2	AR: 0 deg ECEA: 5 deg RR: – 15 deg CI: 8 deg CA: 45 deg	Same	0.100 in.	1¾ in.	0.0075 in. per tooth	150 fpm 260 fpm	390 in. 90 in.	0.012 in.	None
END MILLING — HIGH SPEED STEEL TOOLS										
SAE 4340 Q & T-49 R_c	T-15	35 deg right hand helix CA: 45 deg × 0.060 in. Per. Cl.: 6 deg	¾ in. diameter, 4-flute end mill	0.250 in.	¾ in.	0.001 in. per tooth	55 fpm 80 fpm	70 in. 30 in.	0.016 in.	Soluble oil flood (20:1)
END MILLING — CARBIDE TOOLS										
SAE 4340 Q & T-52 R_c	C-2	AR: 0 deg ECEA: 3 deg RR: 0 deg CI: 15 deg CA: 45 deg × 0.030 in.	1¼ in. diameter, 4-flute brazed tip end mill	0.250 in.	1¼ in.	0.0015 in. per tooth	50 fpm 80 fpm	78 in. 40 in.	0.016 in.	Chemical emulsion* (40:1)
H11 Q & T-52 R_c	C-2	Same	Same	0.250 in.	1¼ in.	0.0015 in. per tooth	60 fpm 150 fpm	105 in. 55 in.	0.016 in.	Soluble oil* (20:1)

Q & T = Quenched and tempered
R_c = Rockwell C hardness
SR = Side rake
BR = Back rake
AR = Axial rake
RR = Radial rake
SCEA = Side cutting edge angle
ECEA = End cutting edge angle
CA = Corner angle
CI = Clearance
Per. Cl. = Peripheral clearance
* Applied as spray mist through axis of cutter.

Ultra Strength Alloys

Material	Tool Material	Tool Geometry	Cutter	Depth of Cut	Width of Cut	Feed	Cutting Speed	Tool Life	Wear-land	Cutting Fluid
SLOTTING — CARBIDE TOOLS										
SAE 4340 Q & T-52 R_c	C-2	AR: – 5 deg ECEA: 1 deg RR: – 10 deg CI: 8 deg CA: 45 deg x 0.030 in.	6 in. diameter, 6-tooth brazed tooth slotting cutter	0.250 in.	1 in.	0.005 in. per tooth	190 fpm 275 fpm	225 in. 150 in.	0.016 in.	None
H11 Q & T-52 R_c	C-2	Same	Same	0.250 in.	1 in.	0.005 in. per tooth	190 fpm 350 fpm	300 in. 180 in.	0.012 in.	None
DRILLING — HIGH SPEED STEEL TOOLS										
SAE 4340 Q & T-50 R_c	T-15	2-flute, 118 deg crankshaft point 7 deg clearance	¼ in. diameter stub drill	0.500 in. through hole	—	0.001 in. per rev.	30 fpm 60 fpm	100 holes 20 holes	0.016 in.	Highly sulphurized oil + light machine oil (1:1)
SAE 4340 Q & T-52 R_c	T-15	Same	Same	0.500 in. through hole	—	0.001 in. per rev.	20 fpm 60 fpm	34 holes 8 holes	0.016 in.	Same
H11 Q & T-52 R_c	T-15	Same	Same	0.500 in. through hole	—	0.0005 in. per rev.	20 fpm 60 fpm	26 holes 7 holes	0.016 in.	Same
AM-350 Solution treated and aged	T-15	Same	Same	0.500 in. through hole	—	0.002 in. per rev.	20 fpm 60 fpm	107 holes 28 holes	0.016 in.	Same
A-286 Solution treated and aged	T-15	2-flute, 118 deg crankshaft point 7 deg clearance	Same	0.500 in. through hole	—	0.005 in. per rev.	20 fpm 60 fpm	51 holes 17 holes	0.016 in.	Same
TAPPING — HIGH SPEED STEEL TOOLS										
SAE 4340 Q & T-50 R_c	M-10	4-flute 60% thread	5/16-18 NC taper tap	0.500 in. through hole	—	—	5 fpm 20 fpm	146 holes 30 holes	Tap break-age	Highly chlorinated oil
	M-10	4-flute 75% thread	Same	Same	—	—	5 fpm	13 holes	Tap break-age	Same
SAE 4340 Q & T-52 R_c	M-10	4-flute 60% thread	Same	Same	—	—	5 fpm	75 holes	Tap break-age	Inhibited trichloroethane
	M-10	Same	Same	Same	—	—	5 fpm	7 holes	Tap break-age	Chlorinated oil + inhibited trichloroethane (3:1)
	M-10	4-flute 60% thread 0.015 in. land	Same	Same	—	—	5 fpm	18 holes	Tap break-age	Same
AM-350 Solution treated and aged	M-10	4-flute 75% thread	Same	Same	—	—	5 fpm 9 fpm	200 + holes 25 holes	Tap break-age	Highly chlorinated oil

Q & T = Quenched and tempered
R_c = Rockwell C hardness
SR = Side rake
BR = Back rake
AR = Axial rake
RR = Radial rake
SCEA = Side cutting edge angle
ECEA = End cutting edge angle
CA = Corner angle
CI = Clearance

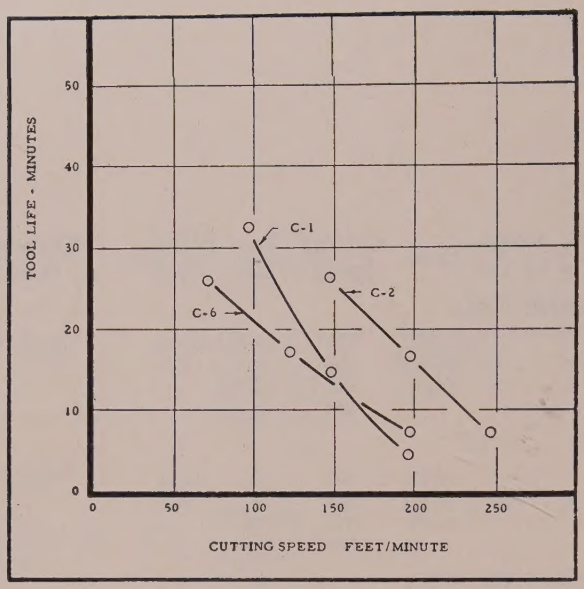


Fig. 7 — Turning A-286, solution treated and aged to 314 Bhn (33 Rockwell C). Tool: carbide (see graph); side rake: 5 deg neg.; back rake: 5 deg neg.; nose radius: 0.032 in.; side cutting edge angle: 15 deg; end cutting edge angle: 15 deg; relief: 5 deg. Mechanical chip breaker. Feed: 0.009 in. per rev.; depth: 0.100 in.; cutting fluid: none; wearland: 0.015 in.

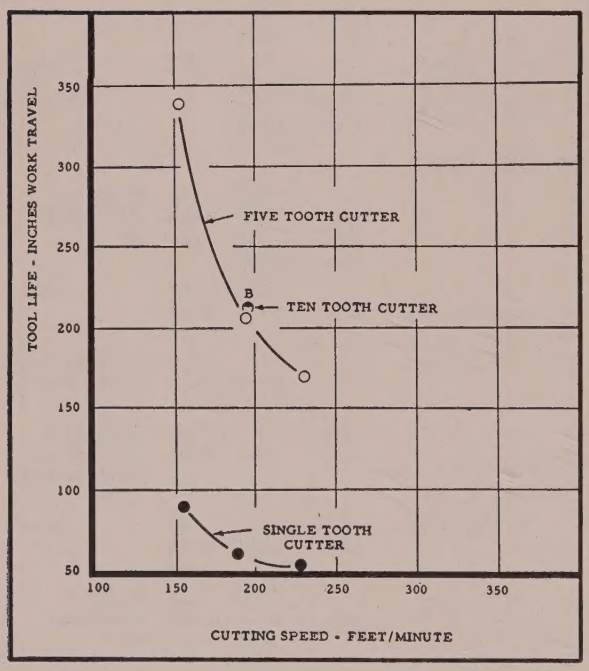


Fig. 8 — Face milling 4340 steel quenched and tempered to 514 Bhn (52 Rockwell C). Tool: 5 in. diameter face mill with C-6 carbide; axial rake: 0 deg; radial rake: 15 deg neg.; corner angle: 45 deg; clearance: 8 deg; true rake: 10 deg neg.; angle of inclination: 10 deg; end cutting edge angle: 5 deg; feed per tooth: 0.005 in.; depth: 0.100 in.; width: 2.25 in.; cutting fluid: none; wearland: 0.015 in.; end of test determined by 0.030 in. localized breakdown.

(Figs. 1–3).

2. SAE designation H11 quenched and tempered to 52 R_c (Figs. 4–5).

3. AM-350 solution treated and aged to 47 R_c (Fig. 6).

4. A-286 solution treated and aged to 33 R_c (Fig. 7).

Table 2 lists recommended carbide grades, tool geometries, cutting speeds, and feeds for the alloys listed above.

Milling Tests

The problem which stands out beyond all others is the milling of the high-strength thermal-resistant alloys in airframe and missile components. The problem was brought about by the change from aluminum to high-strength steel airframe components. Listed below are the important types of milling operations and typical cuts being made for each type of milling operation.

Operation	Dimensions of Cut
Face Milling	1. 0.060–0.250 in. depth 1–5 in. width
End Milling	1. Pockets 4 × 4 in. to 8 × 8 in. area 0.25–3.0 in. depth
	2. Pockets in Thin Sheet 8 × 10 in. area 0.025–0.050 in. depth
	3. Milling Side of Rib 0.020–0.060 in. depth 2–3 in. height of rib
	4. Profiling 0.060–0.300 in. depth 0.5–2.0 in. width
Side Milling	1. Spars 0.5–2.0 in. width ⅛–½ in. depth Up to 25 ft length
Slot Milling	1. Hinges 3/16–½ in. width 0.5–1.0 in. depth 0.5–1.0 in. length
	2. Clevices ¼–1 in. width 1–2 in. depth 1–3 in. length
	3. Wing Spar 1–2 in. width 1–1.5 in. depth Up to 15 ft length

Emphasis will be given in this paper to information and data obtained in milling of SAE 4340 and H11 in the high hardness ranges.

Face Milling Tests

SAE 4340 Quenched and Tempered to 52 R_c — In face milling tests on 4340 quenched and tempered to 52 R_c, a study was made to determine the optimum cutter geometry and tool angles. To establish the optimum tool angles, a series of tests was run in which the angle of inclination was kept constant while the resultant rake was varied. Another series

of tests was made in which the resultant rake was held constant and the angle of inclination was varied. From these tests it was found that a resultant rake of −10 deg and an angle of inclination of 10 deg produced by an axial rake of 0 deg and a −15 deg radial rake with a 45 deg corner angle provided the best tool life when face milling the hardened 4340 steel.

The tool life in face milling is given in inches of work travel. The tool life end point was arbitrarily taken as 0.016 in. uniform wear on the peripheral flank of the cutter tooth, or localized breakdown, whichever occurred first. A 2 in. width of cut and 0.100 in. depth of cut was kept constant for all face milling tests.

With a 5-tooth face mill tipped with a C-6 grade carbide, a cutting speed of 150 fpm and a feed of 0.005 in. per tooth, it was possible to mill 330 in. of material. (See Fig. 8.) At a cutting speed of 190 fpm and a feed of 0.005 in. per revolution, tool life for a 10-tooth cutter was essentially the same as for a 5-tooth cutter. Localized breakdown of the cutter teeth occurred when using the 10-tooth face milling cutter.

Face milling tests made to study the effect of feed rate indicated that a feed of 0.005 in. per tooth was the best rate to use in face milling 52 R_c 4340 steel. The tool life obtained at 0.005 in. per tooth feed was more than twice that obtained when face milling with feeds of 0.0025 in. per tooth and 0.0075 in. per tooth.

Face milling tests on 4340 quenched and tempered to 52 R_c indicated that cutter run-out becomes a very critical factor in tool life. Using a 5-tooth face mill feeding at 0.005 in. per tooth, a cutter run-out of 0.003 in., reduced tool life by about 25% and localized wear developed rather than uniform wear on the cutter. With a total run-out of 0.005 in., tool life was reduced by some 65% and severe localized wear caused cutter breakdown.

H11 Quenched and Tempered to 50 R_c. Cutter geometry studies made in face milling H11 quenched and tempered to 50 R_c indicated that an axial rake of 0 deg and a radial rake of −15 deg, with a 45 deg corner angle, the same as that found to be optimum on 4340 steel, produced the best tool life. The carbide evaluation proved interesting in that a C-2, general purpose, nonferrous grade of carbide produced considerably better tool life than a C-6, general purpose, steel cutting grade.

The tool life curve obtained when face milling 50 R_c H11 shows that with a 5-tooth cutter, a C-2 grade of carbide, and a feed of 0.005 in. per tooth, 400 in. of material can be milled at a cutting speed of 125 fpm. For the same cutting conditions, tool life was only 75 in. when face milling with a C-6 grade of carbide. (See Fig. 9.)

Tests made to study the effect of feed indicated

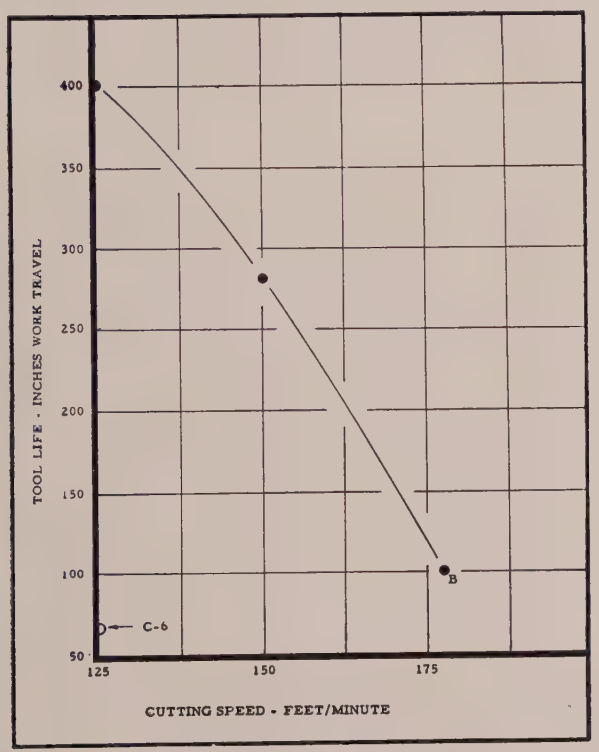


Fig. 9 — Face milling H11 quenched and tempered to 50 Rockwell C; multiple tooth cutter. Tool: 5 in. diameter face mill — 5 teeth with C-2 carbide. Axial rake: 0 deg; radial rake: 15 deg neg.; corner angle: 45 deg; true rake: 10 deg neg.; angle of inclination: 10 deg; end cutting edge angle: 5 deg; clearance: 8 deg; feed per tooth: 0.005 in.; depth: 0.100 in.; width: 2 in.; cutting fluid: none; wearland: 0.015 in. End of test determined by localized breakdown.

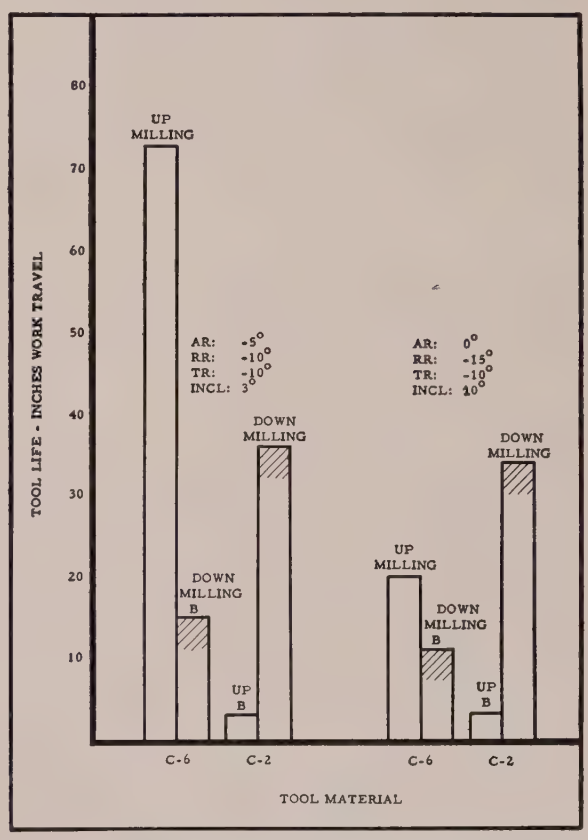


Fig. 10 — Side milling 4340 steel quenched and tempered to 52 Rockwell C; effect of tool material and geometry — single tooth cutter. Tool: 7 in. diameter inserted tooth cutter; corner angle: 45 deg; end cutting edge angle: 5 deg; peripheral clearance: 8 deg; cutting speed: 145 fpm; feed per tooth: 0.0075 in.; depth: 0.100 in.; width: 1.750 in.; cutting fluid: none; wearland 0.015 in. End of test determined by localized breakdown.

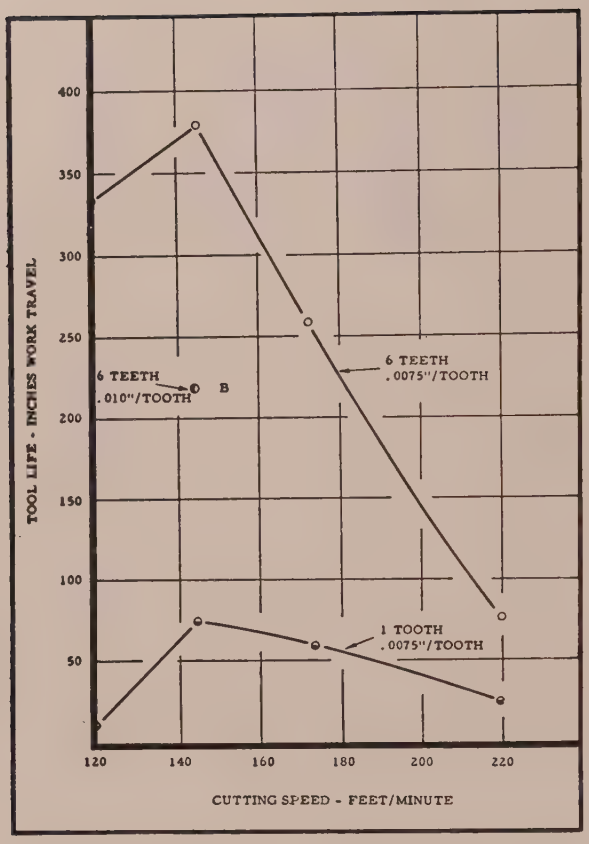


Fig. 11 — Side milling 4340 steel quenched and tempered to 52 Rockwell C; multiple tooth cutter — up milling. Tool: 7 in. diameter inserted tooth cutter with C-6 carbide; axial rake: 5 deg neg.; radial rake: 10 deg neg.; corner angle: 45 deg; true rake: 10 deg neg.; angle of inclination: 3 deg neg.; end cutting edge angle: 4 deg; peripheral clearance: 8 deg; feed per tooth: see graph; depth: 0.100 in.; width: 1.750 in.; cutting fluid: none; wearland: 0.015 in. End of test determined by localized breakdown.

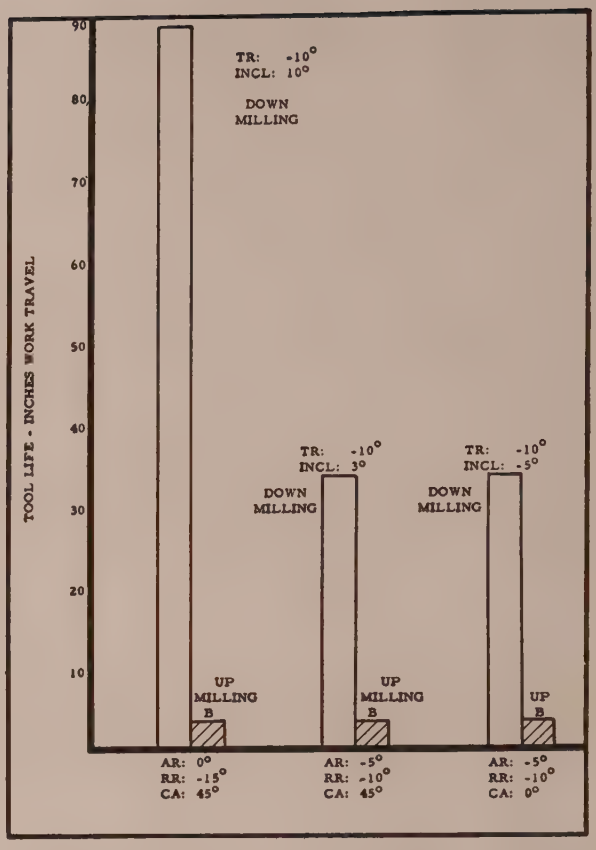


Fig. 12 — Side milling H11 quenched and tempered to 52 Rockwell C; effect of tool geometry — single tooth cutter. Tool: 7 in. diameter face mill with C-2 carbide; clearance: 8 deg; end cutting edge angle: 5 deg; cutting speed: 145 fpm; feed per tooth: 0.0075 in.; depth: 0.100 in.; width: 1.75 in.; cutting fluid: none; wearland: 0.012 in. End of test determined by localized breakdown.

that a feed of 0.005 in. per tooth was best when face milling the H11 alloy. At higher feed rates, from 0.010 to 0.015 in. per tooth, localized rather than uniform wear takes place on the cutter teeth.

Side Milling

SAE 4340, Quenched and Tempered to 52 R_c — Side milling tests were made, using a 7 in. diameter inserted tooth face mill, on SAE 4340 quenched and tempered to 52 R_c. A tool material and tool geometry evaluation using a *single* tooth in the cutter, showed best results could be obtained when up milling, using a C-6 carbide grade cutter with – 5 deg axial rake and – 10 deg radial rake. As shown in Fig. 10, tool life was 72 in. work travel when milling at a cutting speed of 145 fpm and a feed of 0.0075 in. per tooth. When down milling with a C-6 grade carbide, severe tooth chippage was encountered and tool life was only 20 in. of work travel. Down milling was found to be best when using a C-2 grade of carbide; however, maximum tool life was only 50% of that obtained from up milling with a C-6 carbide grade.

The tool life curve (Fig. 11) shows that 380 in. were cut with *6 teeth* in the cutter at a cutting speed of 145 fpm and a feed of 0.0075 in. per tooth. When the feed was increased to 0.010 in. per tooth, tool life decreased to 220 in. work travel.

H11, Quenched and Tempered to 52 R_c — For side milling, H11 quenched and tempered to 52 R_c, down milling with a C-2 grade of carbide was found to give best results (Fig. 12). At a cutting speed of 145 fpm and 0.0075 in. per tooth feed, a tool life of 90 in. length of cut was obtained with a single tooth cutter with 0 deg axial rake and – 15 deg radial rake. Almost immediate cutter breakdown was encountered when attempting up milling at the same cutting conditions. Fig. 13 shows a tool life curve for a 6-tooth cutter. Maximum tool life was 390 in. work travel to obtain a 0.012 in. uniform wearland on all teeth.

End Milling

The end milling tests consisted of milling ¼ in. deep slots using ¾ in. diameter, high-speed end mills and 1¼ in. diameter carbide end mills. Tool life is given in inches of work travel for a tool life end point of 0.016 in. uniform wear or localized breakdown, whichever occurred first.

SAE 4340 Quenched and Tempered to 49–54 R_c — Preliminary tests were made in end milling SAE 4340 quenched and tempered to 49 to 54 R_c with high-speed steel cutters. These tests were designed to establish cutting speeds and feeds that would re-

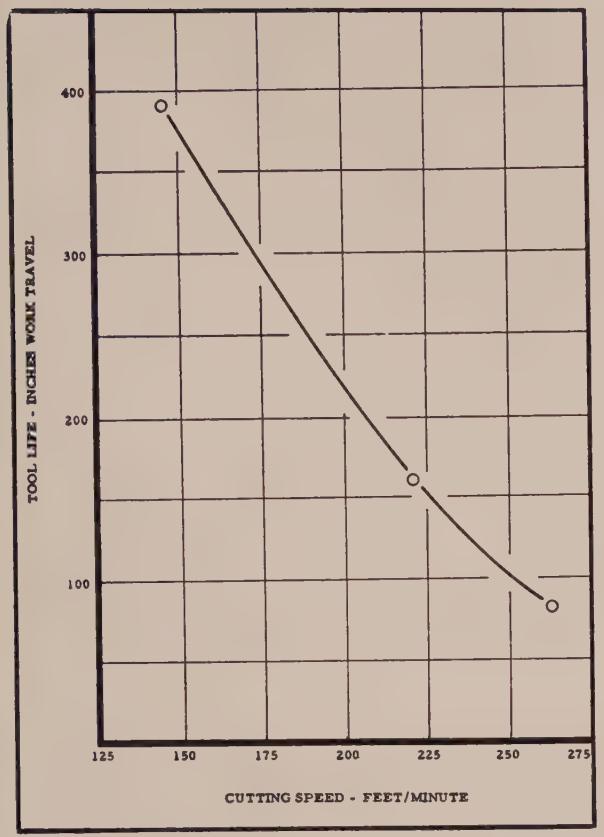


Fig. 13 — Side milling H11 quenched and tempered to 52 Rockwell C; multiple tooth cutter. Tool: 7 in. diameter face mill — 6 teeth with C-2 carbide; axial rake: 0 deg; radial rake: 15 deg neg.; corner angle: 45 deg; true rake: 10 deg neg.; angle of inclination: 10 deg; end cutting edge angle: 5 deg; clearance: 8 deg; feed per tooth: 0.0075 in.; depth: 0.100 in.; width: 1.75 in.; cutting fluid: none; wearland: 0.012 in. End of test determined by localized breakdown.

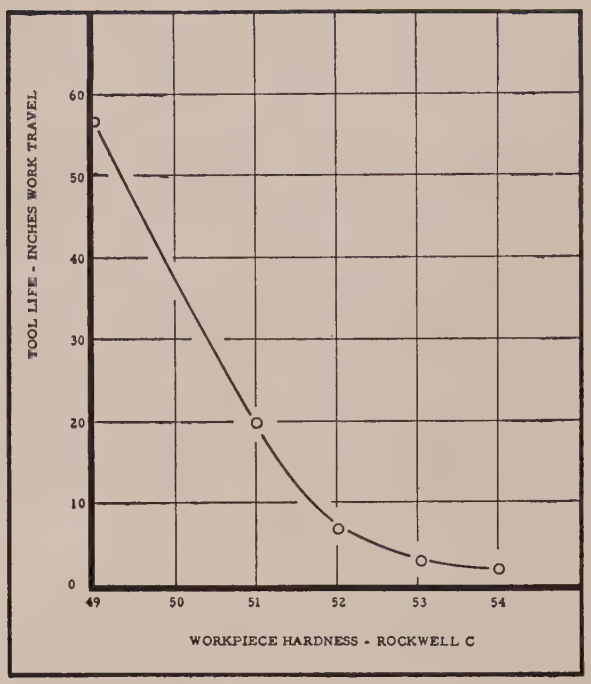


Fig. 14 — End milling 4340 steel — effect of workpiece hardness. Tool: T-15 HSS ¾ in. diameter 4-flute end mill; helix angle: 35 deg right-hand helix, right-hand cut; radial rake: 15 deg; corner angle: 45 deg, 0.060 in. wide; side clearance: 5 deg primary, 10 deg secondary; end clearance: 5 deg primary, 20 deg secondary; speed: 37 fpm; feed per tooth: 0.0005 in.; depth: 0.250 in.; width: 0.750 in.; cutting fluid: soluble oil 1: 20); wearland: 0.016 in.

sult in reasonable tool life. Very inconsistent tool life results were observed, despite careful control of cutter geometry and emphasis on rigidity of the setup. A systematic analysis of all of the data indicated that hardness was a critical variable. A series of tests was then made to determine the effect of workpiece hardness on tool life.

Fig. 14 shows the effect of workpiece hardness on tool life for 4340 steel at hardnesses of 49–54 R_c, using T-15 high-speed steel cutters, a feed of 0.0005 in. per tooth and a cutting speed of 37 fpm. Good tool life was obtained with a workpiece hardness of 49–50 R_c. However, tool life decreased rapidly as hardness increased from 49 to 52 R_c. For workpiece hardness above 52 R_c, the material was, for all practical purposes, unmachinable with the high-speed steel end mills used in the tests.

The tool life curves shown in Fig. 15 for end milling 49 R_c 4340 steel indicate the importance of selecting the correct cutting speed to attain maximum tool life. Using a feed of 0.001 in. per tooth and depth of cut of ¼ in., tool life using a T-15 cutter increased rapidly with decreasing cutting speed between 85 and 55 fpm, then decreased rapidly. For the M-7 cutter, tool life increased rapidly for decreasing cutting speeds between 70 and 45 fpm, then decreased with decreasing cutting speeds.

Tests made to determine optimum feed rates in-

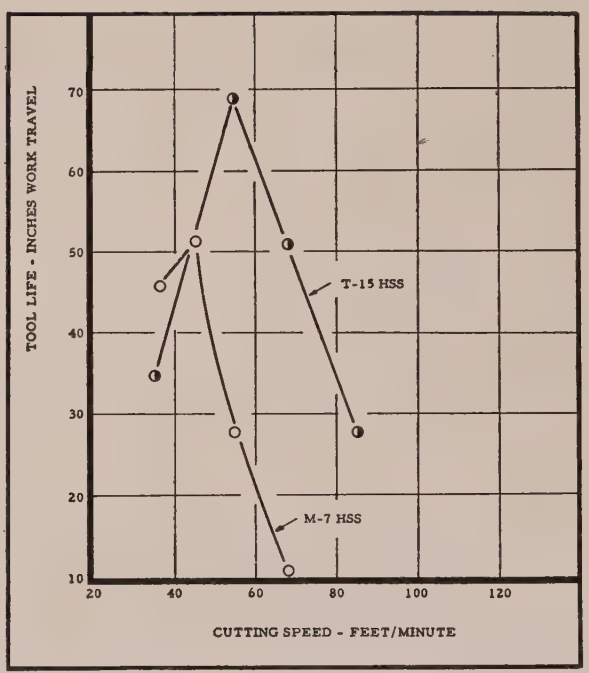


Fig. 15 — End milling 4340 steel quenched and tempered to 470 Bhn (49 Rockwell C). Tool: HSS ¾ in. diameter 4-flute end mill; helix angle: 35 deg right-hand helix, right-hand cut; radial rake: 15 deg; corner angle: 45 deg, 0.060 in. wide; side clearance: 5 deg primary, 10 deg secondary; end clearance: 5 deg primary, 20 deg secondary; feed per tooth; 0.001 in.; depth: 0.250 in.; width: 0.750 in.; cutting fluid: soluble oil (1:20); wearland: 0.016 in.

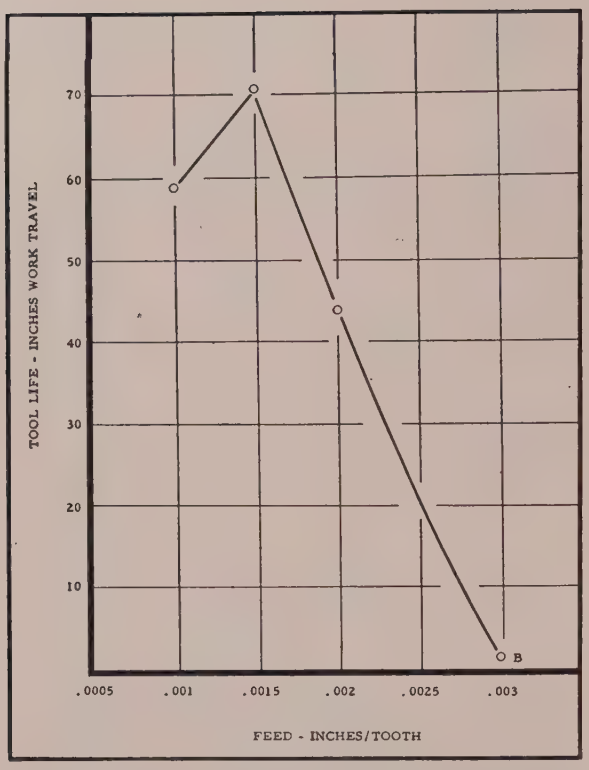


Fig. 16 — End milling 4340 steel quenched and tempered to 52 Rockwell C. Tool: 1¼ in. diameter 4-tooth end mill with C-2 carbide; axial rake: 5 deg neg.; radial rake: 0 deg; true rake: 0 deg; end cutting edge angle: 3 deg; corner angle: 45 deg, 0.030 in. wide; peripheral clearance: 15 deg; end clearance: 5 deg; cutting speed: 50 fpm; depth: 0.250 in.; width: 1.250 in.; cutting fluid: chemical emulsion (40:1) applied as mist through cutter; wearland: 0.015 in. End of test determined by localized breakdown.

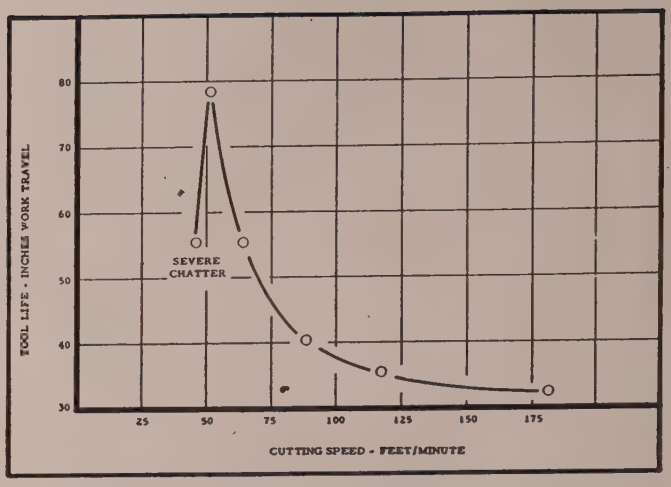


Fig. 17 — Carbide end milling 4340 steel quenched and tempered to 52 Rockwell C. Tool: 1¼ in. diameter heavy duty 4-tooth end mill with C-2 carbide; axial rake: 0 deg; radial rake: 0 deg; true rake: 0 deg; end cutting edge angle: 3 deg; corner angle: 45 deg, 0.030 in. wide; peripheral clearance: 15 deg; end clearance: 5 deg; feed per tooth: 0.0015 in.; depth: 0.250 in.; width: 1.250 in.; cutting fluid: chemical emulsion (40:1) applied as mist through cutter: wearland: 0.015 in.

dicated that feeds between 0.001 in. per tooth and 0.0015 in. per tooth must be used to attain a reasonable tool life in end milling 49 R_c 4340 steel using high-speed steel end mills.

The initial tests in carbide end milling of SAE 4340 quenched and tempered to 52 R_c were all made with standard 1¼-in. diameter, 4-tooth carbide tipped end mills having 1 in. diameter shanks, 1¼ in. flute length and overall length of 4¼ in. Considerable cutter deflection and resultant poor life was observed.

To increase rigidity, special end mills were designed and made with a shank diameter of 1¼ in., and the flute length cut down to ¾ in. These end milling cutters were used in all subsequent tests.

An evaluation of carbides for end milling showed that a C-2 grade carbide provided better tool life than a C-6 carbide grade. In addition, wear on the carbide teeth was uniform and showed no evidence of chipping. With a C-6 steel cutting grade of carbide, tool life was about one-sixth of that obtained with a C-2 nonferrous grade carbide.

End milling of 52 R_c 4340 steel with carbide end mills requires the use of cutting fluid to attain maximum tool life.

Investigation of various methods of coolant application showed that by applying the cutting fluid as a mist through the cutter, an appreciable increase in tool life can be obtained. The coolant mist was fed through a hollow draw bolt through the cutter to the tool work interface. The improvement in tool life is attributed partially to the better chip removal effected by the air pressure used to produce the mist. The feed curve shown in Fig. 16 indicates that optimum feed was 0.0015 in. per tooth in carbide end milling of 52 R_c, 4340 steel.

A tool life of 78 in. of work was obtained in carbide end milling of 52 R_c 4340 steel using a feed of 0.0015 in. per revolution and a cutting speed of 50 fpm. (See Fig. 17.)

H11 Quenched and Tempered to 52 R_c — Tool life for end milling H11, quenched and tempered to 52 R_c with high-speed steel cutters was extremely poor. All subsequent tests were made with carbide tipped tools. As shown in Fig. 18, a C-2 grade of carbide and a mist coolant applied through the cutter yielded the best results. Optimum tool geometry for carbide tipped cutters was found to be 0 deg axial rake and 0 deg radial rake. Chatter and low tool life were encountered for both positive and negative rake angles.

Fig. 19 shows a tool life curve for end milling H11, quenched and tempered to 52 R_c with carbide tipped cutters. Maximum tool life was 105 in. work travel at 60 fpm cutting speed and 0.0015 in. per tooth feed. At lower speeds, chatter and tooth chippage occurred. Tool life decreased slowly with increasing speed dropping to 30 in. at 210 fpm.

Slotting Tests

Machining tests using a 6 in. diameter, 1 in. wide slotting cutter were made on SAE 4340 and H11 steels which were quenched and tempered to 52 R_c. The tests consisted of milling a slot 0.250 in. deep by 1 in. wide in the workpiece on a Cincinnati No. 6 High Power Horizontal mill. To obtain maximum rigidity of the setup a 2 in. diameter arbor was used and the slotting cutter kept as close to the spindle nose as possible. Overhang of the overarm was kept at a minimum. Tool life when slotting was recorded as the inches of work travel before cutter breakdown took place.

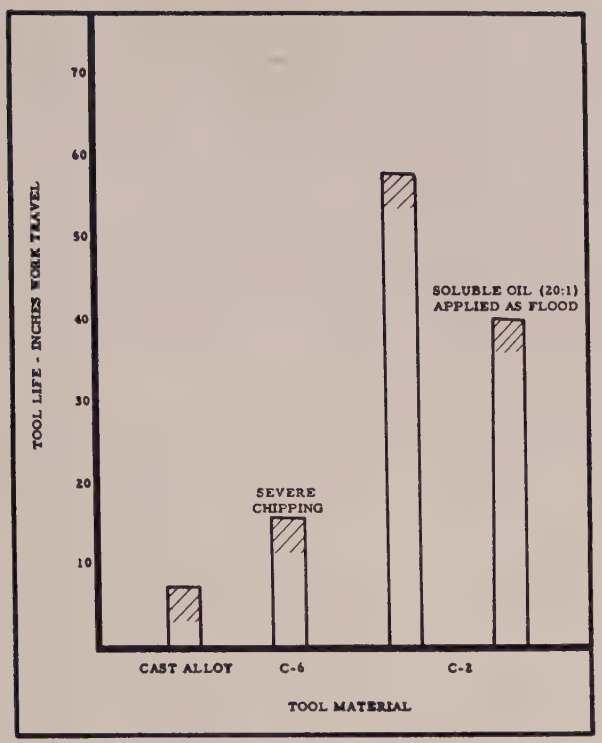


Fig. 18 — End milling H11 quenched and tempered to 52 Rockwell C; carbide cutter — effect of carbide grade. Tool: 1¼ in. diameter 4-tooth end mill. Axial rake: 5 deg neg.; radial rake: 0 deg; true rake: 0 deg; end cutting edge angle: 3 deg; corner angle: 45 deg, 0.030 in. wide; peripheral clearance: 15 deg; end clearance: 5 deg; cutting speed: 50 fpm; feed per tooth: 0.0015 in.; depth: 0.250 in.; width: 1.250 in.; cutting fluid: chemical emulsion (40:1) applied as mist through cutter; wearland: 0.015 in.

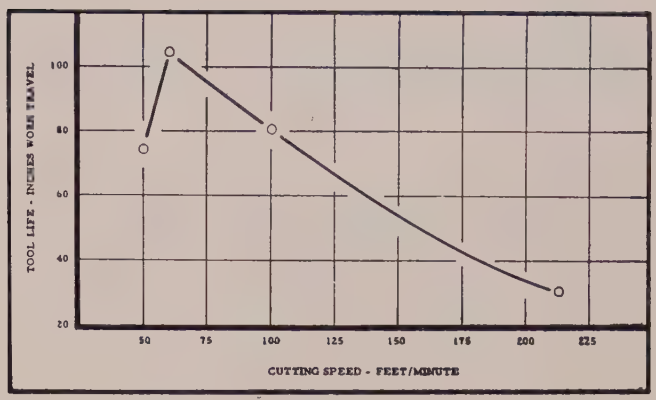


Fig. 19 — End milling H11 quenched and tempered to 52 Rockwell C. Tool: 1¼ in. 4-tooth end mill with C-2 carbide; axial rake: 0 deg; true rake: 0 deg; radial rake: 0 deg; end cutting edge angle: 3 deg; corner angle: 45 deg, 0.030 in. wide; peripheral clearance: 15 deg; end clearance: 5 deg; feed per tooth: 0.0015 in.; depth: 0.250 in.; width: 1.250 in.; cutting fluid soluble oil (20:1) applied as mist through cutter; wearland: 0.015 in.

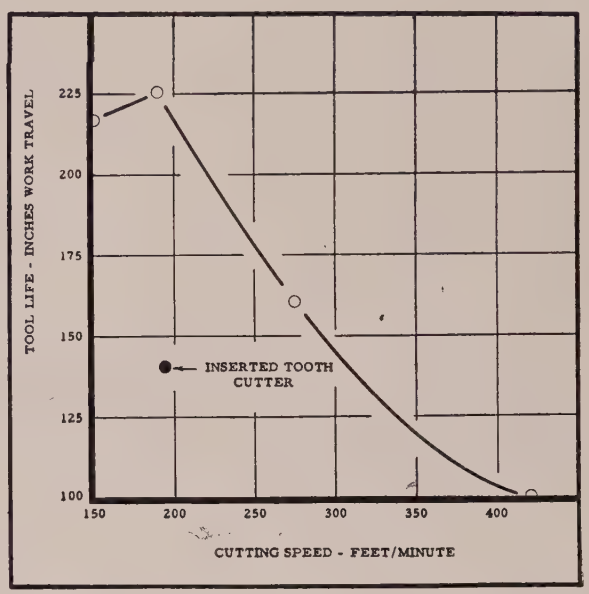


Fig. 20 — Slotting 4340 steel quenched and tempered to 52 Rockwell C; multiple tooth cutter — down milling. Tool: 6 in. diameter 6-tooth cutter with brazed-on C-2 carbide; axial rake: 5 deg bi-neg; radial rake: 10 deg neg.; corner angle: 45 deg, 0.030 in. wide; true rake: 10 deg neg.; angle of inclination: 5 deg neg.; end cutting edge angle: 1 deg; peripheral clearance: 8 deg; feed per tooth: 0.005 in.; depth: 0.250 in.; width 1 in.; cutting fluid: none; wearland: 0.015 in.

The initial testing consisted of a carbide evaluation and rake angle evaluation to determine the best tool material and tool geometry for slotting the hardened alloys. The carbide evaluation showed that the nonferrous C-2 grades of carbide were much better for slot milling than the steel cutting C-6 grades of carbide. With a C-2 carbide tool life was about four times higher than that obtained with a C-6 grade carbide. The C-2 carbide wore very uniformly along the cutting edge, while the C-6 grade fractured and chipped very badly. Tests with a C-3 grade of carbide provided a tool life about the same as that obtained with a C-2 grade, but severe chipping was evident on the teeth.

The rake angle evaluation indicated that a rake combination of bi-negative 5 deg axial rake and negative 10 deg radial rake produced the highest tool life. The face of the carbide tooth was ground to produce the bi-negative 5 deg axial rake on both corners of the tip. When looking at the carbide tip in a radial plane, the cutting edge face shows a wedge shape because of the – 5 deg axial rake ground on both corners. Both corners were ground with a 0.030 in. wide 45 deg corner angle.

Down, or climb, milling shows a 3 to 1 improvement in tool life over up, or conventional, milling in the slotting cuts. With conventional milling considerable vibration was noted in the milling setup even when light feed rates of 0.003–0.005 in. per tooth were used.

SAE 4340 Quenched and Tempered to 52R_c — Fig. 20 shows the tool life curve obtained when slot milling 52 R_c4340 steel. With a 6-tooth cutter it was possible to mill a ¼ in. deep slot 225 in. long at a cutting speed of 190 fpm and a feed of 0.005 in. per tooth. Considerable vibration was encountered when feeds of 0.008 in. per tooth and higher were attempted.

H11 Quenched and Tempered to 52 R_c — The tool life curve when slot milling 52 R_c H11 is shown in Fig. 21. The tool life tests were made using the 6 in. diameter tooth cutter with brazed-on C-2 grade carbide tips. Maximum tool life for a 0.012 in. wear-

land tool life end point was 300 in. of work travel at a cutting speed of 190 fpm and a feed of 0.005 in. per revolution. Good tool life was obtained at cutting speeds from 190–300 fpm.

Drilling Tests

The drilling tests performed on the ultra-strength alloys consisted of drilling ¼ in. diameter by ½ in. deep through holes. Drill life end point was an arbitrary 0.015 in. wearland on the drill margin or complete breakdown, whichever occurred first.

SAE 4340 Quenched and Tempered at 50–55 R_c — Drilling tests made on SAE 4340 steel with hardnesses of 50–55 R_c demonstrated that workpiece hardness was critical with respect to tool life when drilling with high-speed steel drills. (See Fig. 22.) With a T-15 drill, a cutting speed of 30 fpm and a feed of 0.001 in. per revolution drill life was 100 holes for a workpiece hardness of 50 R_c, 30 holes for a workpiece hardness of 52 R_c, and only 5 holes for a workpiece hardness of 55 R_c.

To attain maximum drill life for the hardened steels, it is extremely important to keep drill length at a minimum. Emphasis must also be given to keeping the entire drilling setup as rigid as possible. Results of drilling tests with high-speed steel drills

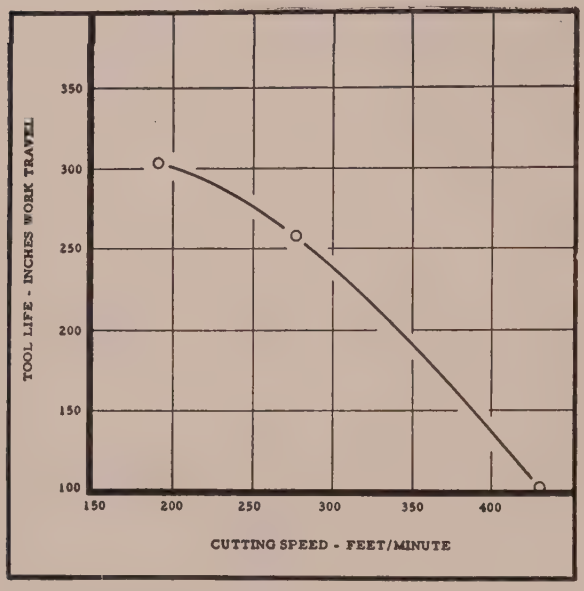


Fig. 21 — Slotting H11 quenched and tempered to 52 Rockwell C; multiple tooth cutter — down milling. Tool: 6 in. diameter tooth cutter with brazed-on C-2 carbide: axial rake: 5 deg; true rake: 10 deg neg.; radial rake: 10 deg neg.; angle of inclination: 10 deg; corner angle: 45 deg, 0.030 in. wide; end cutting edge angle: 1 deg; peripheral clearance: 8 deg; feed per tooth: 0.005 in.; depth: 0.250 in.; width: 1 in.; cutting fluid: none; wearland: 0.012 in.

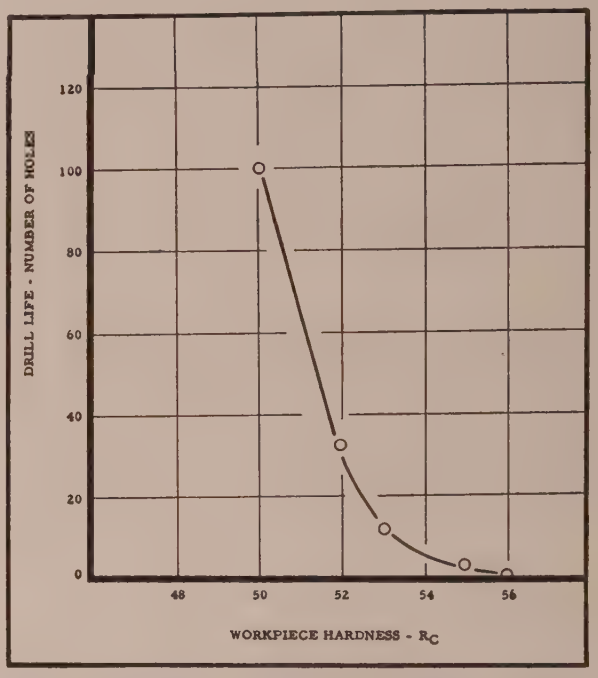


Fig. 22 — Drilling 4340 steel quenched and tempered; effect of workpiece hardness. Drill: T-15 HSS; diameter: 0.250 in.; length: 2.75 in.; clearance: 7 deg; point grind: crankshaft; cutting speed: 30 fpm; feed: 0.001 in. per rev.; depth of hole: 0.500 in.; cutting fluid: highly sulphurized oil diluted 1:1 with light machine oil; wearland: 0.015 in.

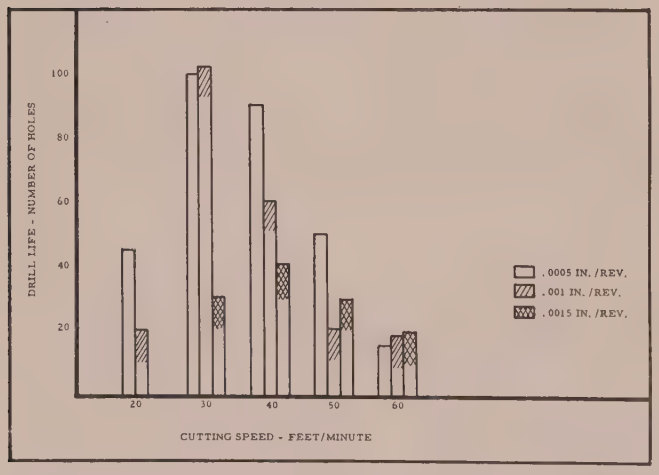


Fig. 23 — Drilling 4340 steel quenched and tempered to 482 Bhn (50 Rockwell C); effect of feed. Drill: T-15 HSS; diameter: 0.250 in.; length: 2.75 in.; point angle: 118 deg; clearance: 7 deg; point grind: crankshaft; feed: see graph; depth of hole: 0.500 in.; cutting fluid: highly sulphurized oil diluted 1:1 with light machine oil; wearland: 0.015 in.

Fig. 24 — Drilling 4340 steel quenched and tempered to 514 Bhn (52 Rockwell C); effect of feed. Drill: T-15 HSS; diameter: 0.250 in.; length: 2.75 in.; point angle: 118 deg; helix angle: 29 deg; point grind: crankshaft; clearance: 7 deg; depth of hole: 0.500 in.; cutting fluid: highly sulphurized oil diluted 1:1 with light machine oil; wearland: 0.015 in.

on 50 R_c and 52 R_c 4340 steel are shown in Figs. 23 and 24 and in the summaries of machining data in Table 2.

H11 Quenched and Tempered to 52 R_c — Drill life data for 52 R_c H11 are presented in Fig. 25 and Table 2.

AM-350 Solution Treated and Aged to 444 Bhn (47 R_c) — A drill life of 107 holes was obtained in drilling 47 R_c AM-350 using a T-15 drill with a crankshaft point, a cutting speed of 20 fpm, a feed of 0.002 in. per revolution, and a highly sulphurized oil. (See Fig. 26.)

Drilling A-286 Solution Treated and Aged to 321 Bhn — Good drill life was obtained in drilling 321 Bhn A-286 using T-15 drills with a crankshaft point, cutting speeds of 20–40 fpm, a feed of 0.005 in. per revolution, and a highly sulphurized oil. (See Fig. 27.)

Tapping Tests

SAE 4340 Quenched and Tempered to 50–52 R_c — Fig. 28 shows that a good tap life of 146 holes was obtained in tapping 50 R_c 4340 using a standard M-10 4-flute taper tap, a highly chlorinated oil, a cutting speed of 5 fpm, and 60% thread. However, for the same cutting condition, tap life decreased

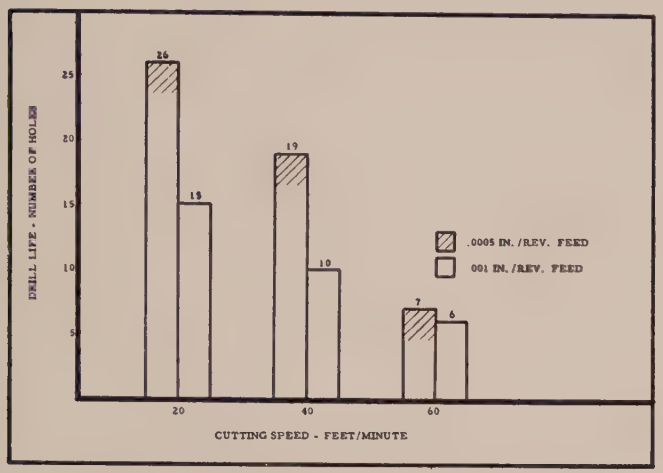


Fig. 25 — Drilling H11 quenched and tempered to 52 Rockwell C; effect of cutting speed and feed. Drill: T-15 HSS; diameter: 0.250 in.; length 2.75 in.; point angle: 118 deg; helix angle: 29 deg; point grind: crankshaft; clearance: 7 deg; depth of hole: 0.500 in.; cutting fluid: highly sulphurized oil diluted 1:1 with light machine oil; wearland: 0.015 in.

Fig. 27 — Drilling A-286 solution treated and aged to 321 Bhn; effect of tool material. Drill: see graph; diameter: 0.250 in.; length: 4 in.; point angle 118 deg; helix angle: 29 deg; point grind: crankshaft; clearance: 7 deg; depth of hole: 0.500 in.; cutting fluid: highly sulphurized oil diluted 1:1 with light machine oil; feed 0.005 in. per rev.; wearland 0.015 in.

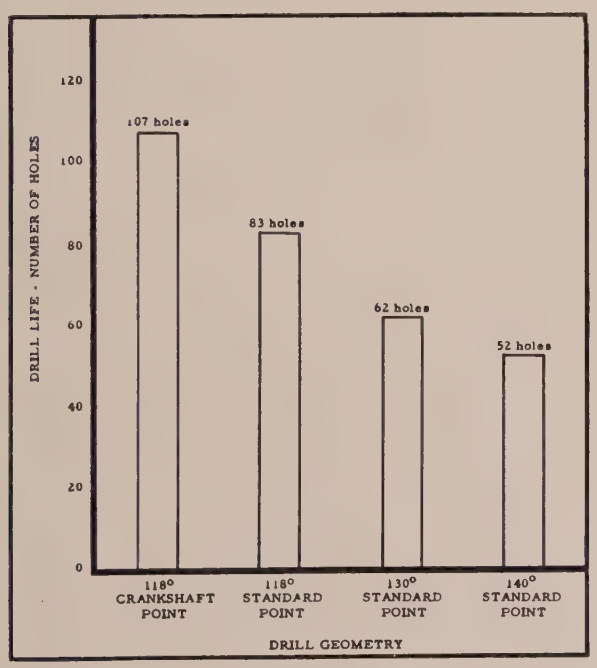


Fig. 26 — Drilling AM-350 solution treated and aged to 444 Bhn (47 Rockwell C); effect of drill geometry. Drill: T-15 HSS; diameter: 0.250 in.; length: 4 in.; helix angle: 29 deg; clearance: 7 deg; cutting speed: 20 fpm; feed 0.002 in. per rev.; depth of hole: 0.500 in.; cutting fluid: highly sulphurized oil diluted 1:1 with light machine oil; wearland: 0.015 in.

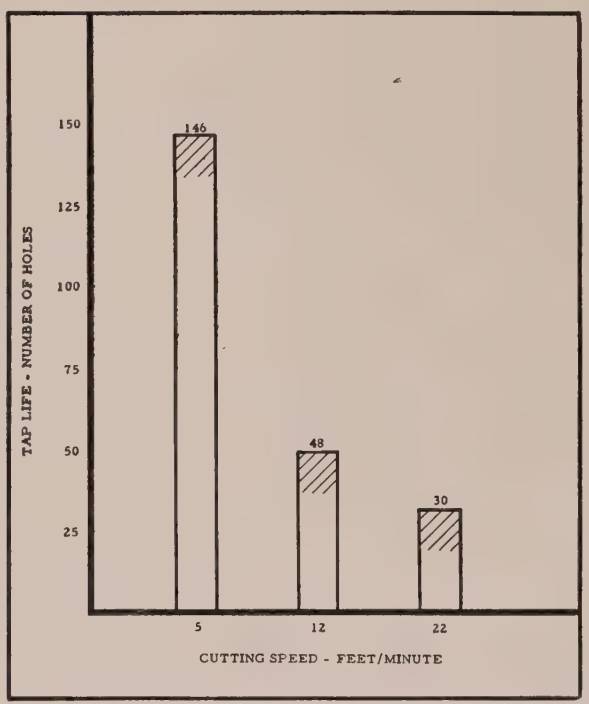


Fig. 28 — Tapping 4340 steel quenched and tempered to 50 Rockwell C. Tap: 5/16-18 NC 4-flute taper; tap material: M-10 HSS; per cent thread: 60%; depth of hole: 0.500 in.; cutting fluid: highly chlorinated oil; end point: tap breakage.

markedly to 18 holes for a 70% thread, and only 13 holes for a 75% thread. (See Fig. 29.)

Tapping of 52 R_c 4340 steel proved to be considerably more difficult than tapping 50 R_c 4340. Using a standard M-10, 4-flute taper tap, a cutting speed of 5 fpm, and 60% thread, the best tap life obtained with any of the commercial types of cutting fluid tried was only 8 holes. The use of inhibited 1, 1, 1 trichloroethane as a cutting fluid resulted in an outstanding increase in tap life to 75 holes. (See Fig. 30.)

Special precautions should be taken when using inhibited 1, 1, 1 trichloroethane as a cutting fluid since it is volatile and toxic. The maximum allowable concentration in air of this solvent vapor per million parts of air is 500, as compared to 25 parts per million of carbon tetrachloride.[1]

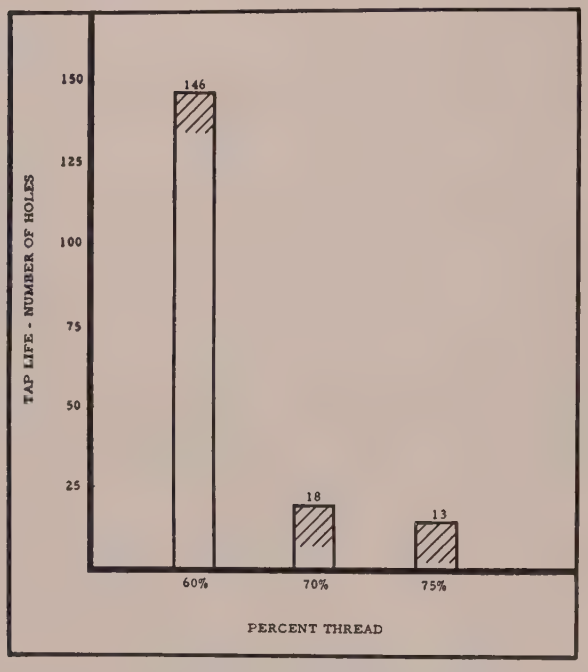


Fig. 29 — Tapping 4340 steel quenched and tempered to 50 Rockwell C; effect of per cent thread. Tap: 5/16-18 NC, M-10 HSS 4-flute taper; cutting speed: 5 fpm; depth: 0.500 in.; cutting fluid: highly chlorinated oil: end point: tap breakage.

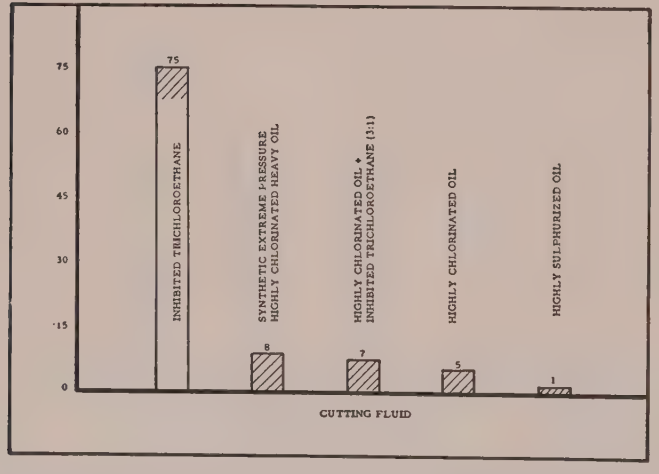


Fig. 30 — Tapping 4340 steel quenched and tempered to 52 Rockwell C; effect of cutting fluid. Tap: 5/16-18 NC 4-flute taper; tap material: M-10 HSS; cutting speed: 5 fpm; per cent thread: 60%; depth of hole: 0.500 in.; end point; tap breakage.

Fig. 31 shows the effect of tap design and surface treatment of taps in tapping 52 R_c 4340 steel. All tests were made using a cutting speed of 5 fpm, 60% thread, and a cutting fluid made up of 3 parts of highly chlorinated oil and 1 part of inhibited 1, 1, 1 trichloroethane. Best tap life of 18 holes was obtained with an M-10, 4 flute taper tap with a 0.015 in. land. Preliminary tests have indicated that tap life can be further improved by nitriding M-10, 4-flute taper taps.

AM-350 — Solution Treated and Aged to 444 Bhn (47 R_c) — Fig. 32 shows that a good tap life of 101 holes was obtained in tapping 47 R_c AM-350 using a standard M-10, 4-flute taper tap, a cutting speed of 9 fpm, 75% thread, and a cutting fluid made up of 3 parts of highly chlorinated oil and 1 part of 1, 1, 1 inhibited trichloroethane.

[1] American Conference of Government Industrial Hygienists, 1957.

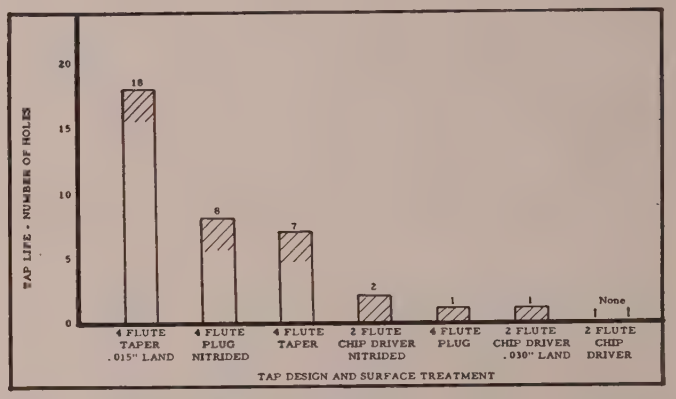


Fig. 31 — Tapping 4340 steel quenched and tempered to 52 Rockwell C; effect of tap design and surface treatment. Tap: 5/16-18 NC; tap material: M-10 HSS — standard and nitrided; cutting speed: 5 fpm; per cent thread: 60%; depth of hole: 0.500 in.; cutting fluid: highly chlorinated oil + inhibited trichloroethane (3:1); end point: tap breakage.

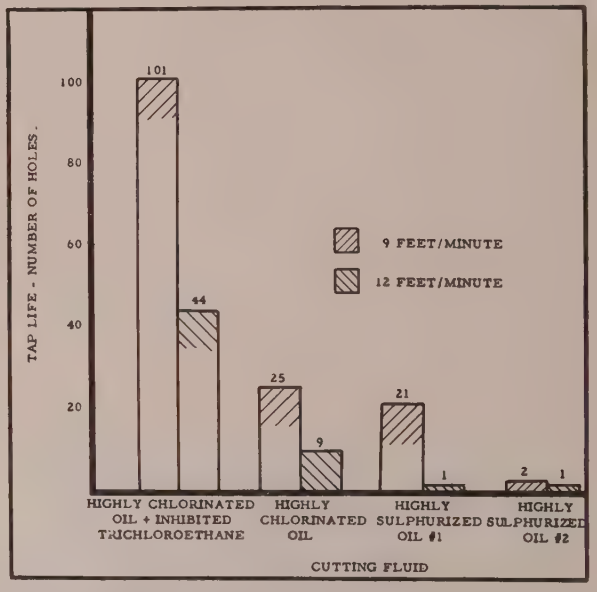


Fig. 32 — Tapping AM-350 solution treated and aged to 444 Bhn (47 Rockwell C); effect of cutting fluid and cutting speed. Tap: 5/16-18 NC 4-flute taper; tap material: M-10 HSS; per cent thread: 75%; depth of hole: 0.500 in.; end point: tap breakage.

Radioisotopes in Research

W. P. Evans and E. W. Landen
Caterpillar Tractor Co

A RECENT report by the National Industrial Conference Board[1] on uses of radioisotopes indicates that the number of industrial laboratories in the United States using isotopes had increased from 18 in 1946 to over 1500 by the middle of 1957. The report also indicates that the annual rate of saving to industry by using isotopes in 1957 amounted to 0.3 billion dollars, and it is expected to rise to 5 billion dollars by 1965. This points out the ever increasing importance of radioisotopes in practical applications in industry and research. In the past decade many industrial companies have acquired and trained personnel to use and handle these radioactive materials safely.

The radioisotope technique is an important research tool. It can provide answers to many problems more quickly and cheaply than conventional methods. In some instances answers are obtained which could not have been determined by any other method. According to the Council for Technological Advancement in Atomic Energy,[2] radioisotopes used in wear tests alone amount to reductions of 75% or more in testing time and 90% in labor and material costs. Not only is time and money saved, but the new technique achieves more accurate results than those obtainable by the older conventional methods.

Radioisotopes are artificially radioactive elements, and they emit nuclear radiations similar to those from the natural radioactive series of uranium and thorium. They can be obtained as off-the-shelf items from the National Laboratories or from qualified commercial laboratories. Proper licensing and safety procedures must be followed. Such regulations are fully described in the *Federal Register*, Title 10.[3] These radioisotopes are produced in the various nuclear reactors and purified by suitable methods into chemical forms adaptable for use in experimental work. Listing of the available radioisotopes may be obtained from the Oak Ridge National Laboratory[4] or from qualified vendors. Radiation can also be induced by irradiation within a reactor. This method is commonly used, for example, in wear studies where the actual piece part, such as a piston ring or a gear, is radioactivated.

This means that the finished component, be it an engine part, gasket, seal, paint, fuel, or lubricant can be obtained either by synthesis or irradiation with traces of activity inherent in the very elements or molecules of which the component is fabricated. Thus, a true tracer is readily available for almost any situation imaginable. Its presence or performance in a system is indicated by its characteristic signal that it continuously sends. This signal cannot be altered by fire, stress, or chemical reaction. It decays only with time.

An important characteristic of a radioisotope is its "half life." This is the time required for the total

---

[1] "Industrial Uses of Radioisotopes," National Industrial Conference Board. Studies in Business Policy, No. 87, 1958.
[2] "Atomic Energy," Council for Technological Advancement, Report No. 33, Jan. 10, 1958.
[3] "Standards for Protection against Radiation," Title 10, Atomic Energy Chap. 1, Part 20. *Federal Register*, Vol. 22, Jan. 29, 1957.
[4] "Radioisotopes, Special Materials and Services," Oak Ridge National Laboratory, Catalog, January, 1957.

RADIOISOTOPES are becoming of increasing importance to industry and research. Applications are of a broad nature and usually contribute to accuracy, sensitivity, or economy. In some instances problems can be solved that would not be possible by conventional techniques.

This paper describes some metallurgical and mechanical applications of isotopes of practical interest to the tractor and earthmoving industry.*

activity present to reduce by 50%. An exponential equation describes this decay process, and is a function of the decay constant of a particular isotope. Fig. 1 shows the general shape of the decay curve. As was mentioned, the rate of decay depends on the isotope. "Half lives" vary from a few minutes or hours to days, months, and years. Half life is an important consideration in planning a tracer experiment.

Because of the importance of steel and iron to the earthmoving industry, it might be appropriate to take a closer look at the isotopes of iron. It should be remembered that isotopes of a given element are chemically similar, differing in mass weight. There are some radioactive varieties and some stable varieties. Table 1 lists the iron isotopes according to Strominger, et al.,[5] designated by mass weight along with some of their pertinent characteristics. When steel is irradiated, the usable radioactive isotope produced is iron 59. This is a strong gamma emitter and easily measurable. The iron 55 isotope deserves consideration only because of its long half life (2.9 years). This is important in disposal considerations.

Radiation hazards and handling procedures have been discussed in the literature[6,7] and the simplest method of assuring proper techniques is to have qualified people. In fact, in order to obtain significant quantities, a company must have trained personnel and the necessary laboratory facilities.

This paper describes some uses of radioisotopes in research which we believe typical of applications of interest to the tractor and earthmoving industry.

Measurement of Wear

The radioactive tracer technique is a desirable method for studying wear of sliding parts and is especially useful in studying gear wear. This is probably the most sensitive method developed for measuring wear quantitatively. Each laboratory develops and uses a particular technique, depending on the conditions being studied. This section describes our method of evaluating wear on sliding surfaces.

As wear on gear teeth is not uniform along the profile of the tooth, an overall measurement of the radioactive wear debris gives the total amount worn from a gear, but does not delineate the wear along the profile of the tooth. In this test we were interested in evaluating the wear in terms of the conditions prevailing at any point on the gear tooth. Since this could not be achieved on our particular gears, we used a pair of rollers to simulate the operating conditions of the gears. In order to show the similarity between the action of the gear teeth and the mating rollers, refer to Fig. 2. At this particular position of the gear teeth the velocity V_1 of tooth A corresponds to velocity V_1 of the active roller, while the velocity of V_2 of the mating tooth B corresponds to velocity V_2 of the nonactive roller. In this manner, operating conditions of the gear teeth are simulated by rollers.

The wear per unit area, W, can be represented functionally and separately for each tooth or for a roller by the formula:

$$W = f(V_1, V_2 - V_1, S_c) \quad (1)$$

where:

V_1 = Velocity of active roller
V_2 = Velocity of nonactive roller
$V_2 - V_1$ = Relative sliding velocity of one member with respect to the other
S_c = Calculated Herz contact stress

The radius of the gear tooth varies along the active face of the tooth, but our present data have been obtained only from rollers having fixed radii. Roller diameters used in this test were 3.30 and 2.70 in., respectively, with the active face width of 0.180 in.; however, the overall width of the rollers was 0.500 in.

The roller test machine has two parallel shafts mounted on roller bearings. Each set of bearings is housed in a massive casting hinged in such a way that a known load can be applied to the rollers. Each shaft can be driven individually, or the two shafts can be interconnected by phasing gears. This provides great versatility in controlling the operating conditions. A steel case surrounds the rotating parts so that the oil is confined in the system. The oil contains suspended radioactive iron which must be kept in a system or in a storage container for disposal. A removable steel cover of ½ in. thickness could be placed over the entire roller machine case for additional shielding against the gamma radiation. This amount of shielding brought the radiation down to a safe level from the activity level of the rollers used. Extra steel blocks were placed between the hot roller and the counting chamber to reduce the background count. A schematic diagram of the roller test setup is shown in Fig. 3. The oil is pumped through the experimental setup by means of a gear pump at a rate sufficient to keep the wear debris in suspension in the circulating oil. From the pump the oil is passed through a counting chamber containing water-cooled geiger tubes. The oil then flows through a heat exchanger to maintain a constant oil temperature and part of the oil flows to the rollers; the remainder flows into the housing

---

* Paper presented at Earthmoving Conference, SAE Central Illinois Section, Peoria, April 15, 1959.

[5] "Table of Isotopes," by D. Strominger, et al. *Reviews of Modern Physics*, Vol. 30, Part 2, April, 1958, pp. 628–629.

[6] "Radiological Health Handbook," by S. Kinsman. U. S. Department of Commerce, Office of Technical Services, January, 1957.

[7] "Safe Handling of Radioactive Isotopes," National Bureau of Standards Handbook 42. U. S. Department of Commerce, NBS, September, 1949.

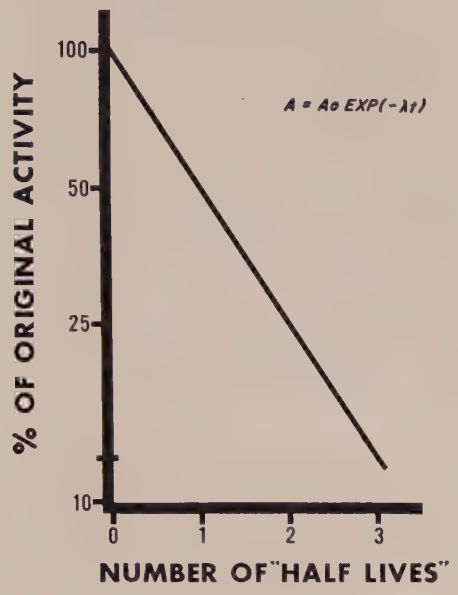


Fig. 1 — Radioactive decay

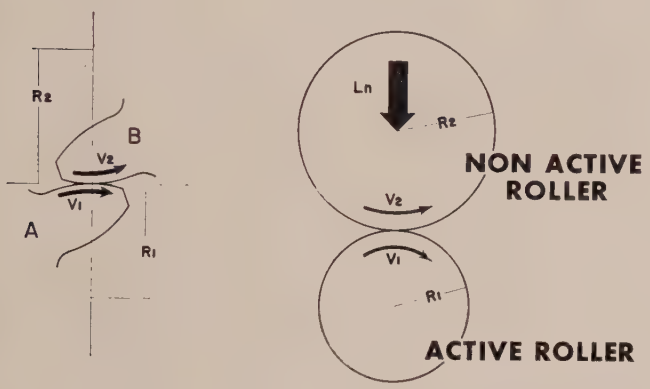


Fig. 2 — Comparison of gear action to rollers

of the roller test machine to keep the wear debris from settling out.

The counting chamber consists of eight side-window geiger tubes in parallel surrounding a pipe in the circulating oil system. The geiger tubes actuate a counter and a chart recorder so that the count rate can be plotted as a function of time.

Calibration solutions of oil-soluble iron were prepared according to the described procedure of Borsoff, et al.[8] Decay corrections were applied from the time that we received the rollers and the initial activity measurements were made.

The initial activity measurements were made in a special setup as shown in Fig. 4. It was found that the rollers were not of equal activity, even though they were activated at the same time. When a package of piece parts is placed in the reactor, there is usually a change in the neutron flux density so

[8] "Tracer Techniques for Studying Gear Wear," by V. N. Borsoff, D. Cook, and J. Otvos. *Nucleonics*, Vol. 10, October, 1952, pp. 67–69.

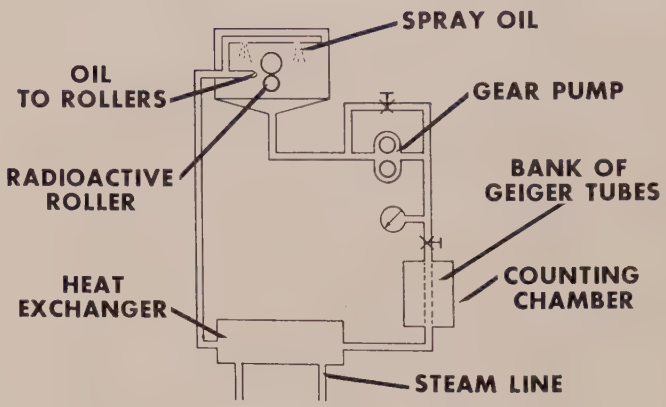


Fig. 3 — Schematic diagram of roller test setup

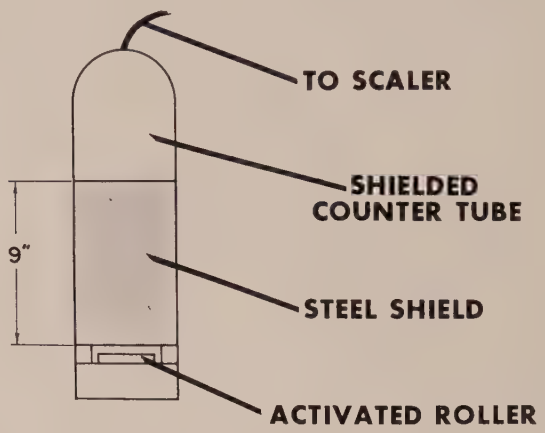


Fig. 4 — Method of measuring initial activity

Table 1 — Iron Isotopes

Stable	Unstable	%	Type Emission	Half Life
	Fe52		β^+, γ	8h
	Fe53		β^+, γ	8.9m
Fe54		5.84		
	Fe55		EC	2.9y
Fe56		91.68		
Fe57		2.17		
Fe58		0.31		
	Fe59		β^-, γ	45d
	Fe60		β^-	3×10^6y
	Fe61		β^-, γ	5.5m

that they are not equally activated. Our initial activity measurements were compared to one of our earliest rollers which we used as a standard. A correction factor was determined and applied to subsequent rollers.

Activity of a piece part placed in a nuclear reactor depends on a number of factors. For a steel roller of 200 grams placed in a neutron flux density of 0.5×10^{12} for 3 weeks, an activity of approximately 20 millicuries of Fe^{59} is produced. One of these rollers when taken from the reactor will produce a gamma intensity of approximately 140 milliroent-

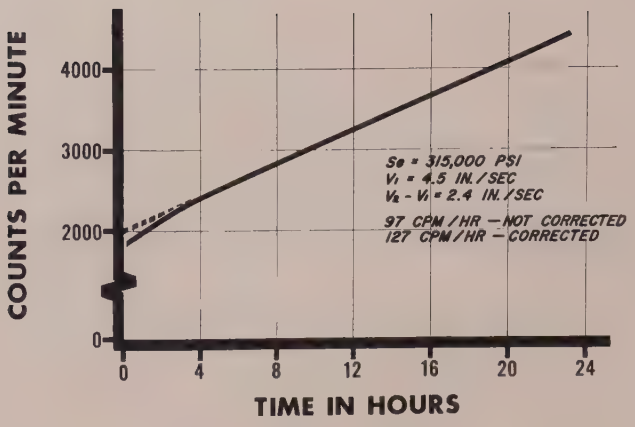


Fig. 5 — Typical raw data curve

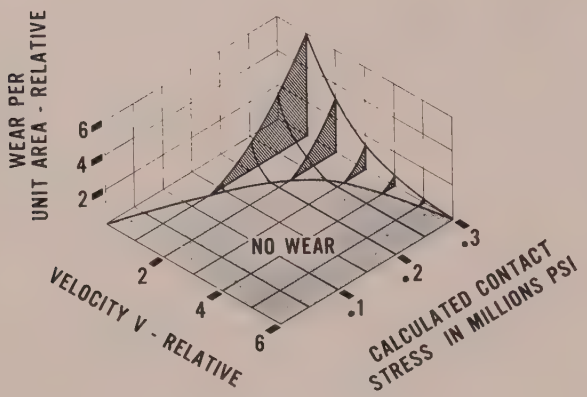


Fig. 6 — Three-dimensional plot of wear

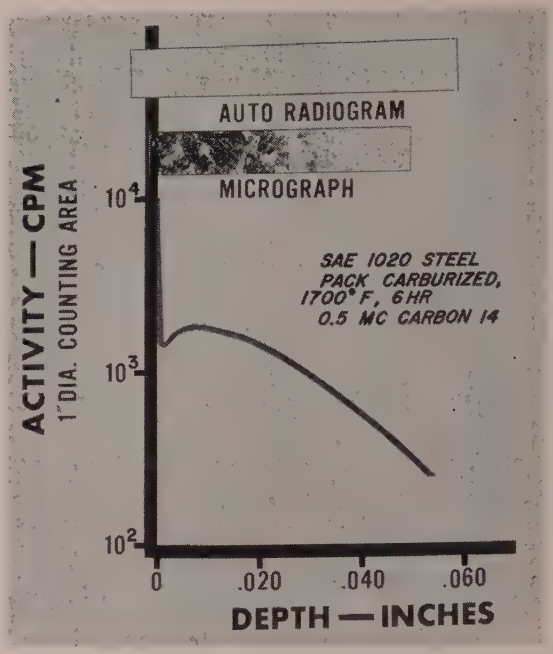


Fig. 7 — Penetration of carbon 14

gens per hr at 1 ft. In planning an experiment of this kind, it is well to keep in mind the available handling equipment, so that you do not get parts that are too hot to handle. Rollers of the above activity can be handled safely, during assembly, with tongs 3 ft long.

A typical graph of raw wear data in terms of activity in counts per minute as a function of time is shown in Fig. 5. Applying the calibration data, the decay correction, and the initial activity correction factor, the slope of this line gives the wear in terms of wear per cycle of the active roller, which can be interpreted in terms of wear per unit area. Each graph of this type gives a wear point depending on the conditions of the test. A number of these points for a given sliding velocity can be represented by an equation, $W = f(V_1 S_c)$ so that a three-dimensional plot is obtained, showing the wear as a function of the velocity V_1 of the active roller and the calculated contact stress S_c. Such a plot is shown in Fig. 6. The shaded area on the velocity-calculated contact stress plane indicates those conditions which produced no measurable wear.

In these tests it was found that a new roller must be used for each run. A used roller which had a wear pattern established did not perform as a newly ground surface. It was also found that even when starting conditions were duplicated as nearly as possible, the same wear rate would not occur. Therefore, a scatter of points was obtained from one operating condition. The scatter of points is probably due to some break-in condition which is an unknown and not controlled. Sufficient tests are, therefore, necessary to obtain proper statistics in analyzing wear data of this type.

Surface Activation

It is of some interest to employ a technique whereby the outer surface layers only of a specimen or part might be radioactivated. This is particularly true when the part is too large to be introduced into a reactor, or so heavy that the induced activity throughout the entire volume would be so high that handling problems would be difficult. The method has application in wear tests where the surface layers only are involved. Surface activation can be accomplished in several ways. Two techniques will be described which are presently being looked into.

The first technique employs the use of carbon 14 as the tracer either in the form of barium carbonate or amorphous carbon. The piece part is pack-carburized in a closed container in the presence of a few millicuries of carbon 14. Diffusion of carbon into the surface layers provides a carburized case containing carbon 14. This method, described by Kohn[9], has been used to good advantage in studying carburizing and other diffusion processes.

Fig. 7 shows the diffused activity of carbon 14 plotted as a function of depth by layer removal technique. The specimens used were 1 in. diameter by 3/8 in. thick steel discs. The layers were removed by grinding the flat faces, and surface activity was measured with a thin mica end-window geiger tube. The accompanying autoradiograph of a section of a

[9] Use of Radioisotopes in Metallurgical Research and Industry," by A. Kohn. *Metallurgical Reviews*, Vol. 3, October, 1958, pp. 158–161.

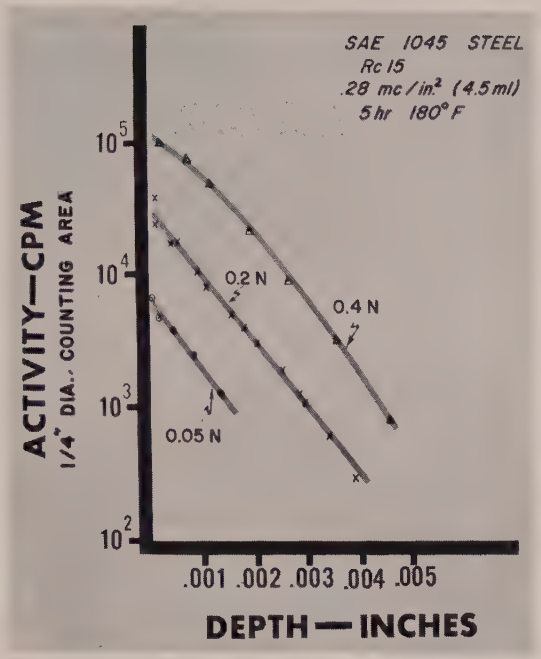


Fig. 8 — Penetration of phosphorus 32

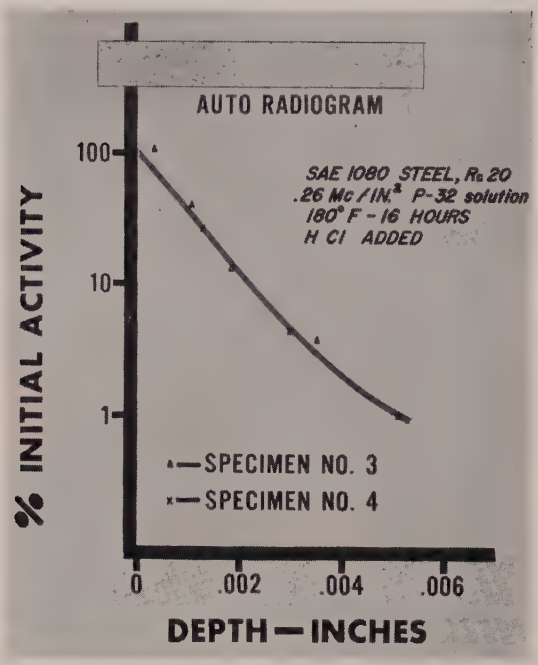


Fig. 9 — Calibration of wear test

similar specimen shows visual evidence of penetration of the isotope. The autoradiograph was made by placing a polished section of the specimen in contact with Eastman Kodak Type NTB-2 emulsion coated on glass plates. As can be seen, it is possible to get activated surface layers 0.010–0.050 in. deep by this method.

The second technique involves soaking of a metallic part in a radioactive solution. Jaoul[10] found that phosphorus 32 in the form of phosphoric acid solution can diffuse into a metallic surface to depths of 10 to 50 microns, depending on time and temperature. We have verified this work in part and have succeeded in reaching depths of penetration of the tracer to 0.003–0.005 in. in 1–8 hr immersion at 180 F.

The mechanism of penetration is not understood. It is too fast for diffusion and according to Jaoul is temperature sensitive in the range of 100–200 F. These temperatures are too low for a diffusion process to be significantly effected. We have found that the induced activity of the surface depends upon the normality of the solution and the specific activity of the solution. A small amount of hydrochloric acid enhances the penetration, although some pitting then takes place. We have removed the pits by lapping and grinding and have found appreciable penetration of the tracer extending approximately an order of magnitude below the pitting depth.

Fig. 8 shows a plot of penetration versus depth as a function of normality of the solution. Disc-shaped specimens, ¾ in. diameter by ⅜ in. thick, cut from SAE 1045 cold-rolled steel were used for this work. Normality was adjusted by small additions of HCl. Successive layers were removed from the flat surface by grinding and lapping. Activity measurement at discrete depths was determined with a thin mica end-window geiger tube. The end of the tube was covered with a brass diaphram having a ¼ in. pinhole in the center.

For the acid concentrations shown, time of immersion did not appear to be a significant factor. Evidently, the reaction and penetration occurs quickly, probably during the first hour of immersion. Depth of pitting due to acid attack was approximately 0.0005 in. The acid also dissolved a thin layer uniformly from the sample as indicated by determination of iron content of the solution after the soak. In the figure shown, depths refer to the original surface before the activation process was begun.

Part of the reaction appears to be of chemical nature and results in the formation of a reddish-brown precipitate. The precipitate is highly radioactive and contains the bulk of the isotope used. Only a small percentage penetrates into the surface layers of the specimen. Assuming an exponential equation of penetration, the 0.4 *N* curve of Fig. 8 gives a mean value of approximately 0.01 millicurie of P-32 absorbed by the specimen of area 1.77 sq in.

When the bath reaches the precipitated stage, the solution is nearly of neutral pH. The P-32 is now in an unusable form and very little additional penetration can occur. The bath can be rejuvenated by redissolving the precipitate with HCl.

Phosphorus solutions were used in activation of the wear test specimens described in Fig. 9. An autoradiograph of a section is included for comparison with that obtained with C 14. In contrast to carbon 14, the concentration of P-32 varies markedly with depth. It is this very gradient that makes wear measurement using this technique attractive. Recovery of wear debris is not necessary inasmuch as local measurement of activity can be made directly upon the specimen. The specimens used were cylindrical in shape, approximately ⅜ in. in diameter by 2 in. long. The calibration graph of

[10] "Study of Die Wear by Means of Radioactivated Surfaces," by B. J. Jaoul. ASME Transactions, Vol. 78, 1956, pp. 1135–1139.

Fig. 9 was obtained by layer removal technique of an end of the specimen not subject to wear in the wear test.

A feasibility study of wear measurement by this technique was subsequently carried out in an abrasive wear test. Results of this test are shown in Fig. 10. Activity was measured near the tip and center of the cylinder at four points 90 deg apart. A brass diaphram covering an end-window geiger tube containing a ⅛-in. pinhole was used in scanning specific areas where measurements were desired. Such a technique appears particularly promising in instances where it is difficult to catch the wear particles, or in cases where local measurements may be of greater significance than average wear rates.

Bleed Rate of Powered Metal Bushings

A sintered powdered metal bushing can be impregnated with oil before being placed in service or placed in service using an oil feed line without impregnation. This test was designed to measure the rate at which oil leaked from an impregnated bushing. This is an experimental test in which the radioisotope tracer is applicable, and a test which we have carried out in our Research Laboratory.

A solution of radioactive zinc (Zn 65) chloride obtained from Oak Ridge National Laboratory served as a tracer in this test. The zinc chloride solution neutralized by hydroxide produced a zinc hydroxide which was dissolved in a naphthenic acid solution to form an oil soluble zinc naphthenate. This concentrated solution of active zinc naphthenate when added to an SAE 30 oil produced a tagged oil which was used for these runs. Zinc 65 has a 250-day half life, and the strong gamma radiations emitted were used for actuating the counting system. It also produces weak gamma as well as beta radiations.

Bushings were impregnated with the active zinc naphthenate-oil solution. The bushing was immersed in a beaker of active oil solution and placed in a vacuum jar so the air could be expelled from the bushing. As the pressure returned to the oil it was forced into the bushing. An initial count was made on the bushing and used as a standard for comparing counts made on this bushing at later times during the test. In addition, these measurements served as a measure of the amount of oil going into the bushing during impregnation.

The bushing test machine consisted of a bushing retainer and a rotating shaft arranged in such a manner that a load could be applied to the bushing relative to the rotating shaft. The bushing was bathed externally with a flow of oil maintained at a pressure of 10–20 psi by a gear pump. Any radioactive oil leaking out of the bushing would then be washed into the circulating oil systems. An electric motor powered the rotating shaft at the desired speed.

A schematic diagram of the oil system and geiger counter tube is shown in Fig. 11. Oil was pumped from the oil reservoir through the counting chamber to the bushing test machine. After the oil flowed past the sintered bushing, it returned to the reservoir. Oil tagged with the soluble radioactive zinc naphthenate would then become mixed with the remainder of the oil and would register as an increase in the count rate from the geiger counter. An increase in count rate indicated the amount of oil which leaked from the bushing as the test continued.

The entire oil circulating system was calibrated by adding small amounts of the oil tagged with a known amount of zinc naphthenate solution. A typical calibration curve is shown in Fig. 12. By means of this calibration curve and the count rate from the circulating oil during a run, the amount of tagged oil which leaks from the sintered bushing can be determined quantitatively.

The type of data obtained by this technique is illustrated in Fig. 13. In this case, the bushing operated at a given load for a 6-hr period, during which time the amount of oil leaking out of the bushing was measured. The bushing was removed and measured independently to determine how much oil remained in it. Then it was returned to the bushing test machine and measured for another 6-hr period. This procedure was continued for the duration of the test.

From the amount of oil leaking from the bushing and measurement of the activity remaining in the bushing, it was possible to determine the amount of oil remaining in the bushing as a function of time. This procedure gave two independent measurements of the oil remaining, and either method was found to be satisfactory. A plot of the oil remaining in the bushing, as determined from activity measurements, is shown in Fig. 14. This is a simple and satisfactory method to study the oil action in powdered metal bushings.

Gasket and Seal Evaluation

These tests were carried out on an engine operating in a tractor at our Peoria Proving Grounds. It was desired to evaluate an experimental arrangement of gaskets and seals in the cooling system of the engine, and attempt to determine quantitatively the amount of leakage from the radiator coolant into the crankcase oil as well as the leakage site. In the past, glycol and alcohol tests have been used, but tests of this type are not conclusive; that is, glycol reacts with oil to some extent and makes chemical determination difficult. Alcohol vaporizes due to the heat of the engine and escapes from the crankcase. Water itself does not provide a good test inasmuch as moisture may condense from water vapor contained in the crankcase. In addition to these factors, it is difficult to pinpoint the leakage site after a leak has developed.

In an effort to eliminate some of these difficulties, a radioactive tracer was added to the coolant. Cesium 134 in the form of a chloride was used. This isotope is a strong gamma emitter, has a 2.3-year half life, is water soluble, and has a high boiling point. It is readily available from Oak Ridge. About 10 millicuries of the isotope were used in this test. By adding a known amount of tracer to the coolant, an accurate check for leakage into the crankcase could be made by measurement of activity increase of the crankcase oil. In addition, an opportunity for tracing the path of leakage was possible. If the leak persisted over a period of time it was hoped that particles of the tracer would be entrained along the path of leakage. Engine parts suspected of allowing leaks could then be autoradiographed on X-ray film after engine disassembly.

The test was run for over 500 hr, during which

time pint samples of lubricating oil from the crankcase were periodically removed for radioactivity determination. Engine oil was changed at intervals of approximately 120 hr. Make-up coolant was added as necessary from a make-up barrel containing the same tracer concentration as in the radiator. The result of this part of the investigation is shown in Fig. 15. Apparently, leakage began after about 175 hr of engine operation and continued from then on approximately uniformly at an average rate of 0.5 cc per hr. After completion of the test, various gaskets and seals were evaluated by the previously mentioned autoradiographic technique. Exposures for the autoradiograms varied from 10–20 hr.

Reproductions of sections of the head gasket are shown in Fig. 16. The blackened exposed areas indicate presence of the tracer. The sharp dark lines indicate cracks on the bottom side of the gasket near the water holes, which could also be seen visually. Evidently, leaking and diffusion of coolant was occurring around the water holes near the cracks. The cracks may have occurred after 175 hr of operation,

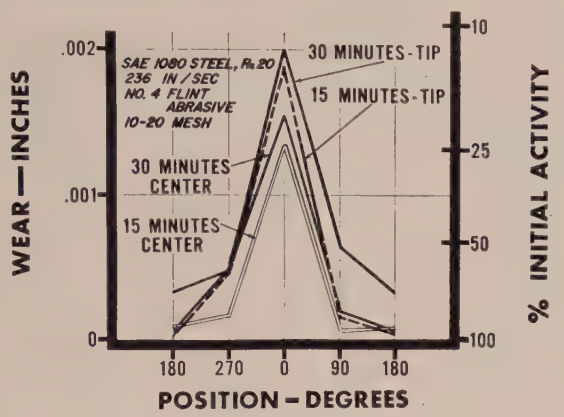


Fig. 10 — Abrasive wear

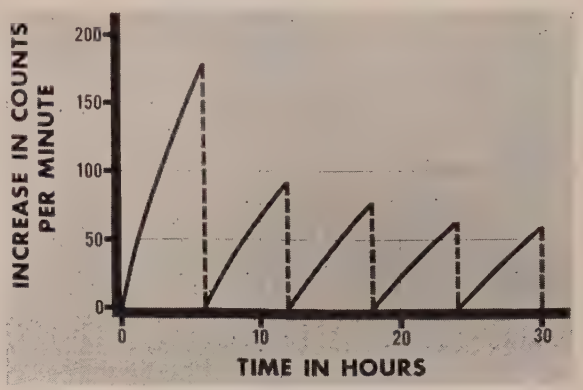


Fig. 13 — Counts per minute from active oil leaching out of bearing during 6-hr runs

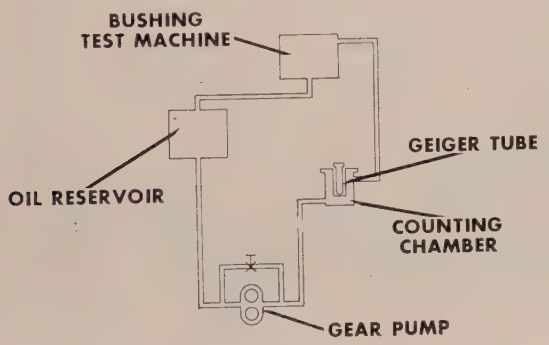


Fig. 11 — Schematic diagram of bushing test setup

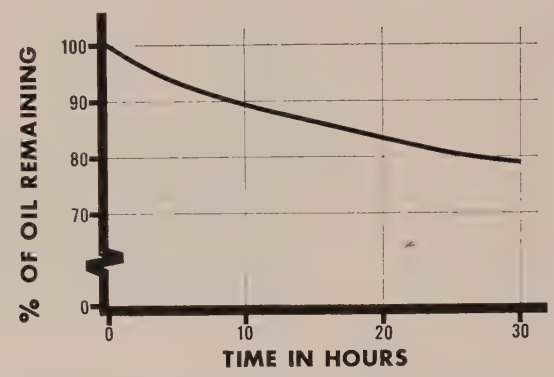


Fig. 14 — Oil remaining in bushing as function of operating time

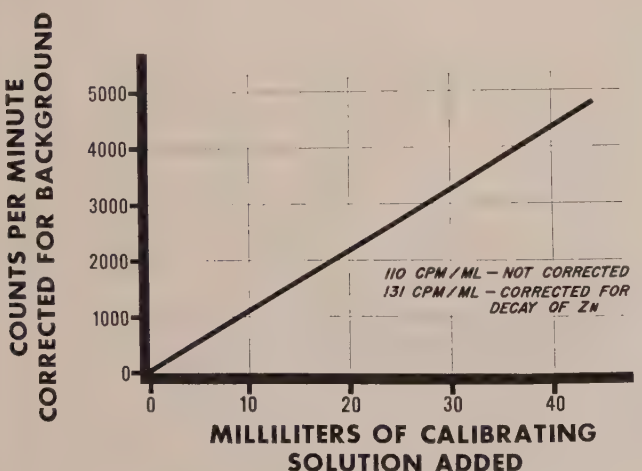


Fig. 12 — Calibration of Zn active oil in bushing test machine

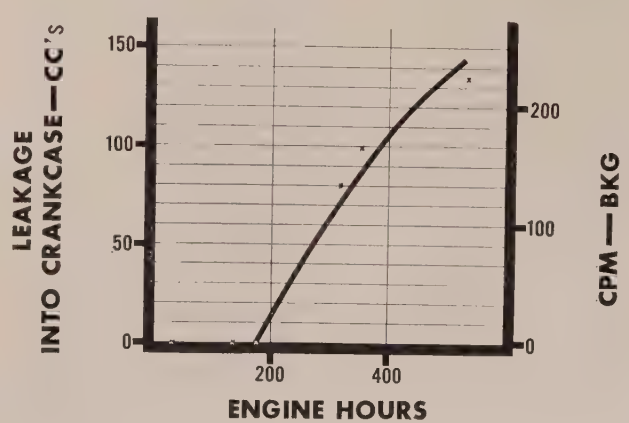


Fig. 15 — Radiator leakage as function of engine hours

which could have initiated the leakage. Collection of coolant under the crimped edge on the top side of the gasket indicates possible path of leakage into the camshaft housing and oil which was adjacent to this side of the gasket. Autoradiograms of the cylinder liner seals indicated good sealing here. There are three "O" ring seals in the bottom of each liner. The top "O" ring is next to the radiator coolant, while the bottom is next to oil from the crankcase. Although there was slight blackening from the center seal on some autoradiograms, this was not considered to be significant. The center seal was almost always dry, indicating good sealing. Autoradiograms of the precombustion chamber gaskets (Fig. 17) indicated that one had leaked. This would have contributed to leakage into the crankcase only when the engine was not running. The tractor involved was run continuously (three-shift operation) except on weekends when it was shut down. Improper sealing of the head gasket was

Fig. 16 — Autoradiogram of opposite sides of cylinder head gasket

Fig. 17 — Autoradiogram of precombustion chamber gaskets and film technique

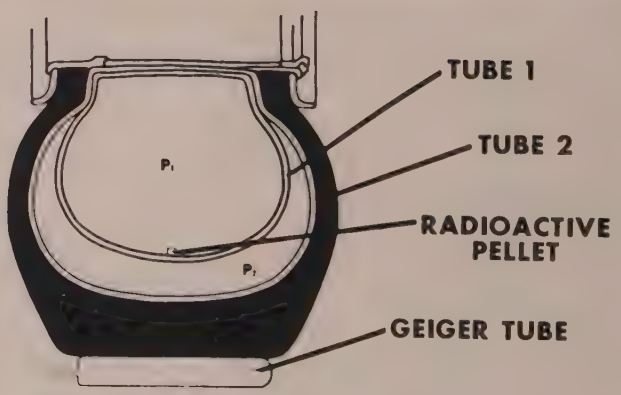


Fig. 18 — Cross-sectional diagram of composite innertube and tire

considered the main cause of leakage into the crankcase during this particular test.

Radioactive Position Indicator

A measurement of an inaccessible place such as the inside of a tire can be made by means of a small radioactive pellet and a geiger counter. This is a simple means of determining the location of the inner member of a double-walled inner tube. This inner tube was made in such a way that there were two compartments which could be individually inflated. Fig. 18 shows a cross-sectional diagram of the double inner tube in a standard tire casing. It will be noted that tube 1 is suspended within tube 2. At the time, we had on hand some small radioactive iron pellets which would drop through the valve stem of the inner member. In this figure the position of the pellet can be seen relative to the geiger counter which was placed below the tire casing. Then, as the pellet approached the geiger tube, the count rate would increase and the position of tube 1 could be located by means of a previously established calibration curve. To illustrate the method of location, the pressure P_2 in the outer compartment was maintained at 45 psi while the pressure P_1 in the inner compartment was varied. Fig. 19 shows the separation of the two members of the inner tube as a function of the pressure P_1 in the inner compartment. In this test an isotope enabled us to get the desired information in a short time.

[11] "Techniques of Industrial Radiography," by G. R. Forrer. Institute of Nuclear Studies, Oak Ridge, Tenn., April 26, 1955.
[12] Obtained from British Steel Castings Association, Surrey, England.

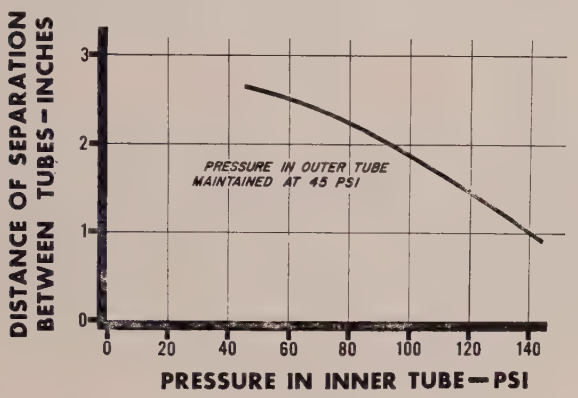


Fig. 19 — Distance of separation as function of pressure

Gamma Radiography

This is an example where the radiation of an isotope is used for its penetrability. It is similar to X-ray in this respect and can be used for inspection of hidden defects or flaws. Techniques of gamma radiography have been previously described in the literature; see for example Forrer.[11] The commonly used isotopes for this work are cobalt 60, cesium 137, and iridium 192. Because of its higher energy, cobalt is used for the thicker sections up to about 6 in. of steel. Iridium gives good results up to about 2 in. of steel and for the less dense nonferrous alloys. Cesium lies in between cobalt and iridium in regard to penetrability and resolution.

Our radiographic facilities are shown in Fig. 20. The walls are concrete blocks, reinforced with steel rods and filled with poured concrete. Exposures are projected toward the back wall, which is 2 ft thick. The side walls are 8 in. thick and absorb the scattered radiation. An electric eye alarm system monitors the entrance. The room is equipped with an overhead crane hoist for handling large parts. The room was designed for 1 curie of cobalt 60 or equivalent. We are presently using 0.3 curies of cobalt and a maximum of 15 curies of iridium.

Exposure times for gamma sources are somewhat longer than those required for corresponding X-ray sources, but economy and mobility of the gamma sources more than compensate for this. The capability of panoramic exposure is another point in favor of the gamma source. Exposure time is easily calculated as a function of film type, thickness of part, film to source distance, and kind of source being used. We use a very convenient slide rule for the purpose.[12] A graph of exposure time versus

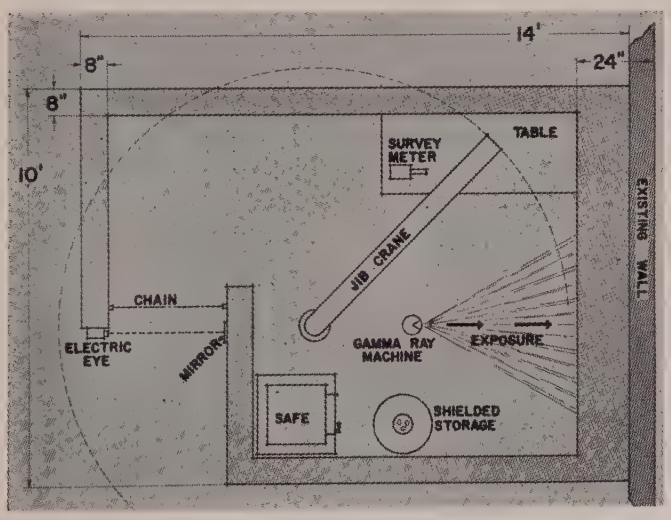


Fig. 20 — Radiography laboratory

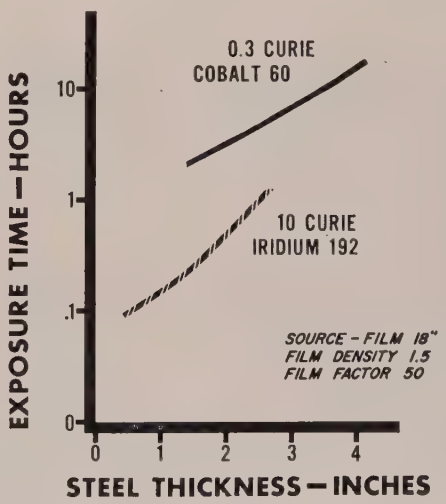


Fig. 21 — Radiographic exposure as function of source

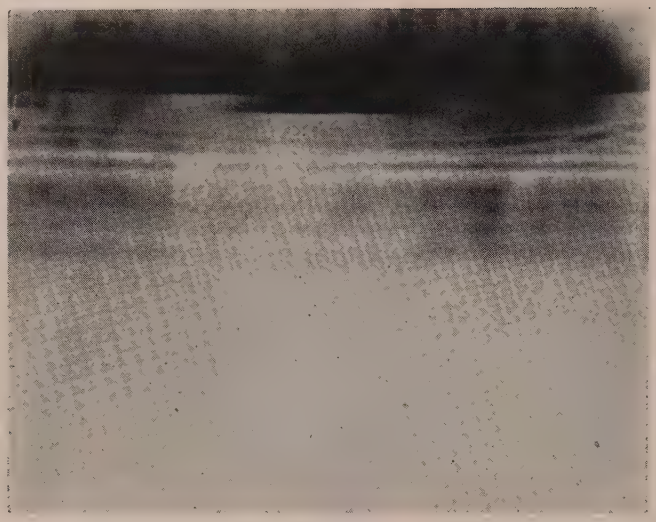


Fig. 22 — Panoramic radiogram of composite piston

Fig. 23 — Radiogram of experimental casting

steel thickness as a function of source is shown in Fig. 21.

As previously mentioned, an important advantage of the gamma source is its versatility. Fig. 22 shows a radiograph of a composite aluminum piston. It was desired to locate the cooling coil with respect to the valve clearance holes and ring grooves for a machining modification. In this case, iridium was used and the source was protruded inside the piston. Exposure time was 45 sec, calculated for about 10 curies of iridium. X-ray film was wrapped around the outside of the piston, giving a panorama of the piston walls. In this way an approximate radial exposure was obtained, which located the coil with respect to the ring grooves and valve clearance holes.

Another use that can be made of radiography is examination of specimens and experimental parts prior to testing. This can be a time and money saver in that flaws and defects can be caught before too much expense has gone into the test. Fig. 23 shows rather marked porosity in an experimental casting. This not only pointed the way toward improved casting technique but also saved money in machining costs and testing time. The radiogram was made with about 7 curies of iridium exposed for 30 min.

The next two examples give an indication of contrast obtainable with gamma sources. Fig. 24 shows a radiogram of a vacuum valve using about 10 curies of iridium exposed for 20 min. Fig. 25 shows a turbine rotor of a type used in turbochargers. The 0.3 curie cobalt source was used with an exposure time of 16 hr.

Conclusion

We have given several examples where we, as manufacturers of earthmoving equipment, have used radioisotopes. We believe there are many cases where builders of heavy machinery of the kind used in earthmoving could employ the technique of radioisotopes as a valuable research and production tool. The use of radioisotopes has increased many fold in the past decade, and many industries are accepting them in research, control, and measurement.

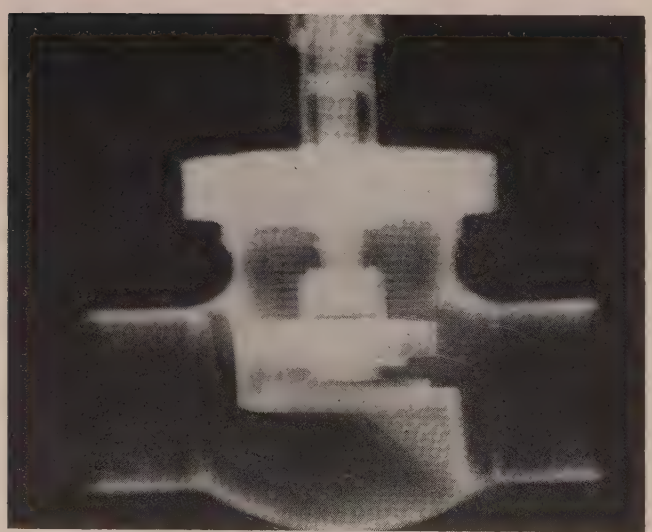


Fig. 24 — Radiogram of vacuum valve

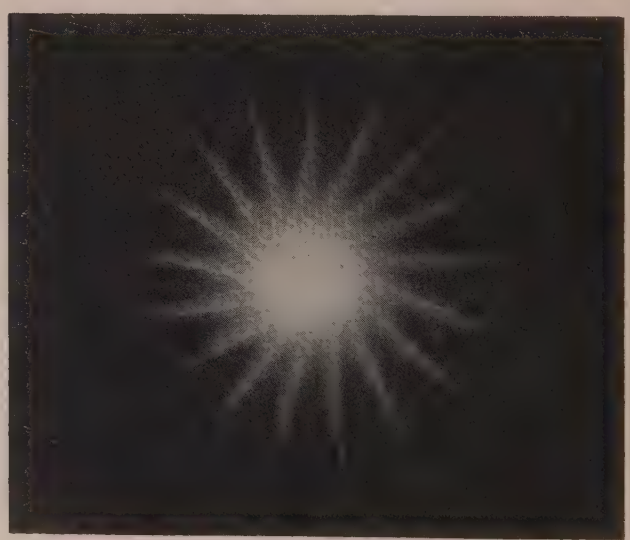


Fig. 25 — Radiogram of turbine rotor

basic design of

Turbochargers

for diesel engines

Hans Egli
AiResearch Industrial Division, The Garrett Corp.

ALTHOUGH turbocharging of low-speed diesel engines has been used world-wide for a long time, it is only during recent years that it has been applied to high-speed diesels. This is the result of considerable engineering efforts from both the turbocharger and the diesel side that were put into the turbocharger, which appears to be a so utterly simple device. This paper describes some of these engineering efforts.

The basic design characteristics are developed with the point of view in mind that the turbocharger has become much more than just an additional accessory. It is a vital component of the basic engine itself, contributing actively to the advancement of this prime mover. The basic design characteristics center heavily around aerodynamical and thermodynamical performance criteria which are so important in any advanced high-speed turbomachine.

In discussing the turbocharger, being a combination of two turbomachines which handle large volume flows of air and gas, emphasis is necessarily put on the compressor and turbine, whereas other components such as bearings and seals are only touched upon briefly.*

THE TURBOCHARGER, when used to its full benefit, has become a very vital and powerful component of the basic diesel engine. It is a self-contained booster device which, without requiring shaft power for its drive, provides the engine with more air for combustion. This enables the engine to burn more fuel and to develop more power within a given displacement. Although air is cheap, it is as important for combustion as expensive fuel. With the exception of the radiator cooling air, the air intake and gas exhaust volume flow rates are the largest ones of all the fluids in the engine. The turbocharger does not need any power for its drive, nor does it produce any shaft power in a direct form. Nevertheless, the turbocharger's own internal power transmitted from its turbine to its compressor is of quite some magnitude and easily reaches 20% of the engine output for today's degrees of turbocharging. With higher boost pressures up to the levels used on free-piston engines, the turbocharger internal power reaches 100% of the diesel power.

In order to write down the basic requirements and design characteristics for a turbocharger a look at the combination of diesel and turbo is in order.

Fig. 1 shows in solid lines a typical indicator diagram for a naturally aspirated 4-stroke engine. It is striking to observe that the major contribution to the indicator diagram output area is in the region of low cylinder volume and high pressures. The portion at the large volume end of the diagram contributes only little to the useful output, however,

* Paper presented at SAE Metropolitan Section Meeting, New York, March 10, 1959.

it determines the size of the engine. The dotted lines correspond to a turbocharged diagram of the same output per pound of charge air (220,000 ft-lb/lb) where the lost portion of the diagram at large volume has been compensated for by an increase in peak pressures. The turbocharged diagram is more "fat" with a corresponding increase in bmep as illustrated by the two rectangular areas.

Whereas the naturally aspirated engine compresses the charge air from practically ambient conditions of specific charge air volume (cu ft/lb) at point *A* to its final compression volume given by the engine compression ratio (a volume or density ratio), in the case of turbocharging this compression is shared by the turbocompressor and the engine. The compressor of the turbocharger takes care of the first portion of the overall compression ratio, from *A* to *B*. Therefore, the density ratio or compression ratio (this is a volume ratio and not a pressure ratio), which is handled outside the engine cylinder, is a very first important design characteristic of the induction system consisting of the turbocharger compressor with or without aftercooler. Its definition is:

$$\text{Density ratio } \sigma = \frac{\text{Air density at intake manifold conditions}}{\text{Air density under ambient conditions}}$$

In a very first approximation the bmep is proportioned to the density ratio unless limited by mechanical loads or heat loads of the engine.

Just as an illustration, Fig. 2 shows how attractive an increase in bmep is from the point of view of engine size. For instance, doubling the bmep will reduce engine bulk (without accessories) by a factor of 2.83 for the same power, number of cylinders, piston speed, and bore-to-stroke ratio.

Of course, it would not make sense to increase the bmep if the added turbocharger (plus in many cases the aftercooler) would occupy too much space, add too much weight, or add too much initial or maintenance costs.

Ingenuity in design of turbochargers and turbocharged engines and effective production methods have taken care of the cost aspects to make turbocharging attractive even to the builder and user of high-speed diesel engines.

Fortunately turbomachines have a very large gulp capacity compared to reciprocating engines, as illustrated by Fig. 3. The mean piston speed of the reciprocating engine, which determines the total pis-

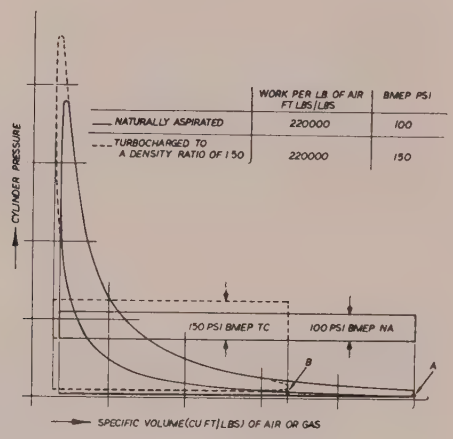


Fig. 1 — Typical diesel indicator diagram

$$BHP = \frac{BMEP \times DISPL \times RPM}{792000} = \frac{BMEP}{132000} \times A_{PISTON} \times V_{PISTON}$$

$$\text{ENGINE VOLUME} = \text{CONSTANT} \times \left(\frac{1}{BMEP}\right)^{3/2}$$

FOR CONSTANT OUTPUT, NUMBER OF CYLINDERS AND MEAN PISTON SPEED.

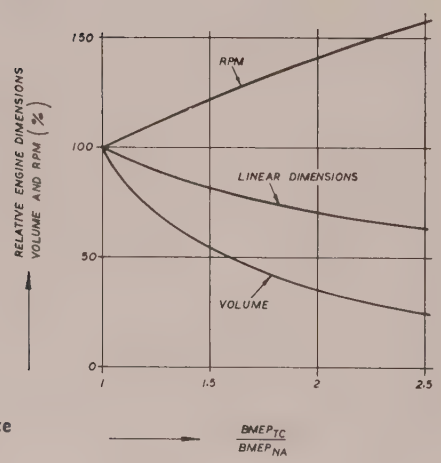


Fig. 2 — Engine size — bmep relation

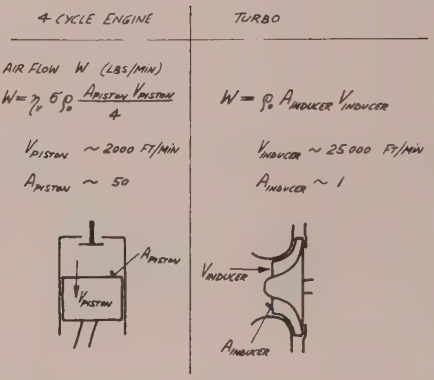


Fig. 3 — Gulp capacity comparison

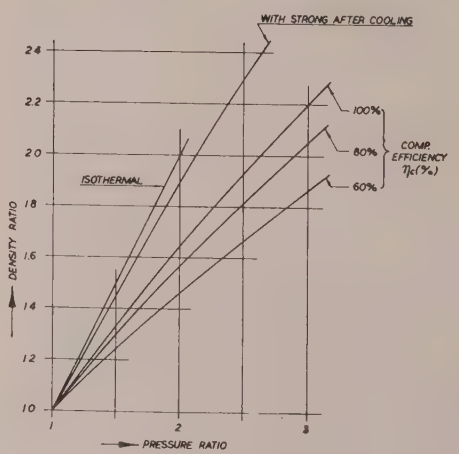


Fig. 4 — Pressure ratio — density ratio relation

ton area and, therefore, the size, is limited for mechanical reasons. Furthermore, a factor four or two shows up for the 4- or 2-stroke engine because only one of the strokes is used for aspirating air. No such reduction factor applies to the turbocompressor because of its continuous and uninterrupted flow. No such mechanical limitations exist for the turbocompressor. Here, the inlet eye or inducer flow area is aerodynamically limited by the velocity of sound of the air. This results in a ratio of characteristic areas (piston area to inducer area) of around 50 for the example shown in Fig. 3.

When talking about compressors, one usually speaks of pressure ratio and not of density ratio. Fig. 4 shows the relationship between these two ratios, which is dependent of the compressor adiabatic efficiency. The engine likes charge air of high density at the lowest possible pressure which clearly points out that high compressor efficiency is very much desirable. The graph also shows the desirability of charge aircooling.

In further clarification, Fig. 5 shows these relationships for a given density ratio of 1.6. It is evident that low compressor efficiency requires high boost pressure (solid line) which also produces a high intake manifold air temperature (dotted line), resulting in high combustion-chamber compression and firing pressures together with a general rise in engine thermal loading.

Fig. 6 (again for a density ratio of 1.6) indicates that high compressor efficiency is not only desirable for the engine itself but also for the turbocharger proper, requiring lower tip speeds of the impeller. Reason for this reduction in tip speed is two-fold: first, because the pressure ratio is smaller; and second, because for any pressure ratio a highly efficient compressor requires a smaller tip speed assuming the same basic compressor blading type. Since the tip speeds of compressors are very high anyway, the resulting reductions in impeller stresses due to centrifugal forces are really significant.

It is often thought that aftercooling eliminates the need for high compressor efficiency, since the aftercooler reduces the high compression end temperature to a low and desirable value, no matter if the compressor has a high or low efficiency. This is only partially true. Even neglecting any considerations of aftercooler size and engine scavenge differential, Fig. 6 shows that compressor efficiency retains its importance because of its influence on impeller tip speed and impeller stress level.

For a selected impeller material and process of

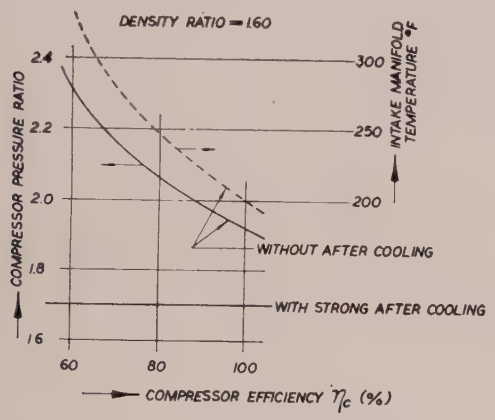


Fig. 5 — Influence of compressor efficiency on intake manifold conditions

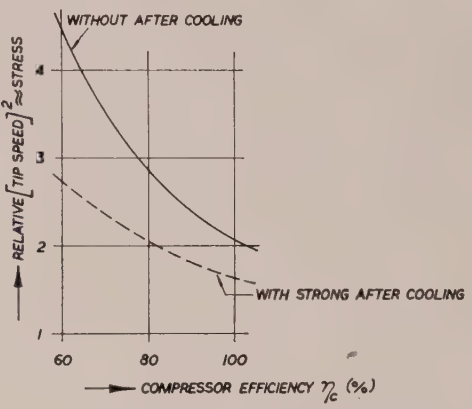


Fig. 6 — Influence of compressor efficiency on impeller stress

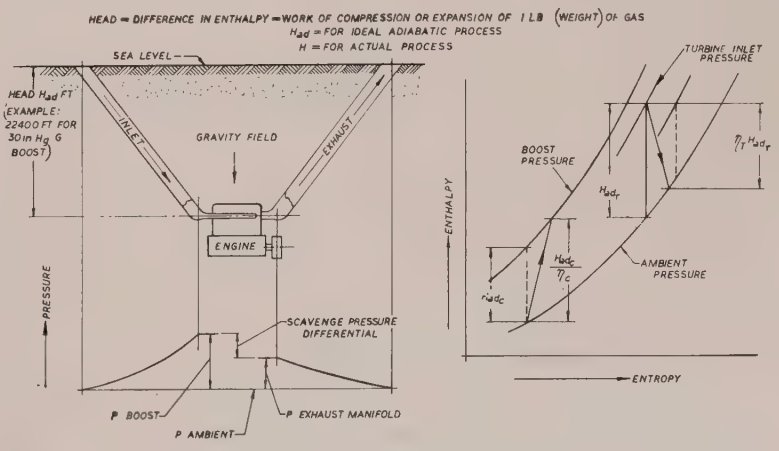


Fig. 7 — "Head" and energy balance

ENERGY BALANCE: $\frac{W_C H_{ad_C}}{\eta_C} = W_T H_{ad_T} \eta_T \eta_M$

manufacturing, there is a definite upper limit of acceptable stress level in the impeller. It is obvious that low compressor efficiency will require an impeller design which is heavier, has a larger moment of inertia, contains more stored kinetic energy of rotation, results in higher bearing loads, and in the case of an overhung rotor arrangement will also increase the overhung rotor moment.

Fig. 7 shows an illustration to the so-called "head" a quantity which is frequently used in turbomachine analysis. The head is the work of compression or expansion per pound weight of working fluid. The picture shows an engine buried deep in the ground. It represents some kind of "turbocharging" process due to the field of gravity. In this case, it is the gravity field which provides the work of compression while air moves down the inlet "stack" and receives back the same amount when the exhaust gas moves up the exhaust stack (neglecting friction losses). The corresponding air and gas pressures are shown in the sketch at the bottom. Due to the different densities on the air and exhaust side, the difference in pressure between the sea level and underground is larger on the intake side than on the exhaust side, resulting in a pressure difference across the engine called scavenge pressure differential. It is obvious that this pressure differential becomes smaller if the process of compression in the intake stack (compressor) and of expansion in the exhaust stack (turbine) is not free of friction. On the right side of Fig. 7, the compression and expansion processes are shown in an enthalpy (head) versus entropy chart: dotted for ideal adiabatic processes; and solid if the two components, compressor, and turbine (expander) are operating at a diabatic efficiencies below 100%. The figure also shows the energy equation which expresses the power balance between the two components, that is, compressor and turbine.

Now to the question of turbocharger overall efficiency, which is defined as the product of compressor, turbine, and mechanical efficiency, as follows:

$$\eta = \eta_C \, \eta_T \, \eta_M \qquad (1)$$

Fig. 8 indicates what a variation in turbocharger overall efficiency means to the engine. The turbocharger contributes to the engine shaft power output also in an indirect fashion by providing a pressure in the exhaust manifold which is below the pressure in the intake manifold, contrary to the conditions on naturally aspirating engines. The figure shows this scavenge pressure differential for three different degrees of turbocharging (boost pressures) and for two different engine exhaust temperatures.

The scavenge pressure differentials reach quite considerable values. They assist the engine and contribute to the indicator diagram output area by making the indicator loop area for the exhaust and intake stroke less negative or even positive. This is how a turbocharger of high overall efficiency contributes in an indirect pneumatic way to an increase in engine shaft horsepower.

A good scavenge pressure differential gives the engine improved breathing capability, improves volumetric efficiency, and further increases the quantity of trapped clean air for combustion. This is the reason why the initial increase in bmep when going from a naturally aspirated engine to a slightly turbocharged engine is larger than the increase in density ratio. It also removes one of the limitations of engine mean piston speed.

One frequent usage of the scavenge differential was to increase combustion-chamber scavenging by increasing the valve overlap. In addition to improving the scavenging of 4-stroke engines (which is not bad to start with) this additional airflow is used for cooling the combustion chamber, piston crown, and exhaust valve. From the point of view of heat transfer, this method of cooling the combustion chamber is definitely not economical. It is, however, very simple. There are many turbocharged engines which use fairly conventional valve overlaps for different reasons.

Even small turbochargers (in the 300-engine-hp class) reach overall efficiencies in the order of magnitude of 60% with single manifolds, and higher with subdivided manifolds for good exhaust pulse utilization. The question arises rightfully: What are these tremendous scavenge differentials good for? Can the 4-stroke engine really make full use of them?

Fig. 8 also leads to the conclusion that high overall efficiency is mainly important at part load (low exhaust temperatures). It appears that at high load one can sacrifice some overall efficiency, which means turbine efficiency, since it has been shown before that compressor efficiency should be high by all means. The turbocharger designers have taken advantage of this fact by building turbochargers which are relatively small by accepting a larger kinetic energy loss at the turbine discharge. This resulted in lightweight turbochargers with a small moment of inertia with a corresponding improvement in response rates during transients.

Another way to utilize the high overall turbocharger efficiency is the application of the pressure ratio control system as built by the AiResearch Industrial Division. This control system is only in operation when the engine is at high loads, that is, when some bypassing of energy can be afforded.

Turbomachines Used for Turbochargers

Fig. 9 shows four different types classified according to their flow pattern in a meridional plane.

During the history of turbochargers their type of compressors and turbines reflected the general state of the art in those fields of turbomachinery. The very first turbochargers had a compressor of the pure radial type (I) with a shrouded impeller, and an axial turbine (IV). Later, due to the great advances made in the industrial and aircraft field a few turbochargers have been built using multistage axial compressors (IV) for experimental purposes. The subsonic axial compressor produces only a small pressure rise per stage, and requires multistaging which makes its usage less attractive from an economical point of view. With turbocharging to higher pressure ratios the pure radial compressor (I) disappeared because of its strength limitations and made place for the radial compressor per II (without or with shroud) and the mixed flow compressor per III. Among the turbines we find the three types II, III, and IV today. Looking into the future, it appears that the type II compressor will continue to lead the field. On the turbine side it is likely that the axial turbine will maintain its field

for the large turbochargers, whereas in the field of small turbos it will probably become increasingly more difficult to make a real distinction between the radial and the mixed flow type.

Radial Compressors

Fig. 10 shows how this type of compressor works. Swirl-free air, that is, air which does not rotate around the axis of the machine, is aspirated and enters the bladed impeller through the inlet eye. The blades impart a swirling motion upon the air which leaves the impeller at its OD. The work which has been put into the air is equal to $\frac{U\,Cu}{g}$ for every pound of air, where U is the impeller tip speed and Cu is the circumferential component of the velocity C of the air at the impeller discharge. It is to be noted that quite a sizeable portion of this work input is in the form of kinetic energy $\frac{C^2}{2g}$ at the impeller exit. Therefore, another part is necessary—the diffuser, which converts this kinetc energy into useful rise of static pressure by slowing the high-velocity air carefully down to an acceptable value.

There are different types of impeller blades in use. The impeller exit velocity triangle at the left of Fig. 10 corresponds to a 90-deg blading, the one at the right to a backward curved blading. Both types of blades have their advantages and disadvantages. In comparison to the 90-deg blading the backward curved blading requires a higher tip speed, the leaving velocity C, however, is smaller which means that a larger percentage of the overall pressure rise takes place within the impeller itself, and the diffusor becomes less important (high degree of reaction). This has two major consequences: (1) the maximum pressure ratio for impellers of similar disc shapes is lower because of strength considerations, and (2) the operating range is wider. In addition, the peak compressor efficiency is usually higher.

Both types of blades have been widely and successfully used in turbochargers. A single turbocharger model using a backward curved bladed impeller can cover a wide range of engines of different airflows for medium pressure ratio turbocharging. This is quite significant when it comes to questions of total production volume, spare part stock, and maintenance. The 90-deg compressor turbocharger will require more turbocharger models to cover the same range of engines. It will, however, permit a larger power boost.

The simplest type of diffusor design is a so-called vaneless diffusor. It does a good job in conjunction with backward curved impeller blades for compressors requiring a wide operating range. Vaned diffusors are usually used together with 90-deg compressors. Because the air leaves a 90-deg impeller with very high velocity, approaching or even exceeding the velocity of sound, vaned diffusors are often combined with short vaneless diffusors in order to improve the operating range.

The small graph at the bottom of Fig. 10 illustrates the pressure rise, the losses in the impeller and in the diffusor, the total work input, and the

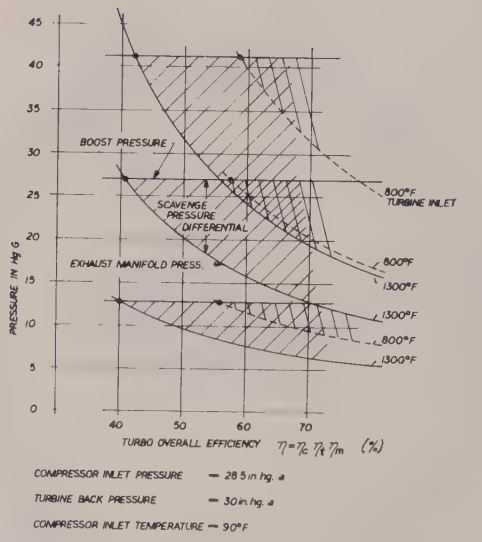


Fig. 8 — Engine scavenge pressure differential

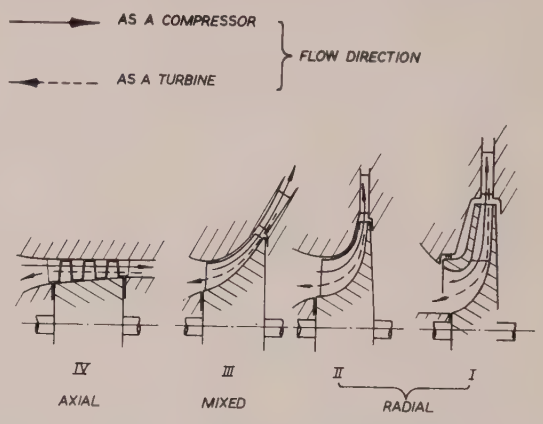


Fig. 9 — Types of turbomachines

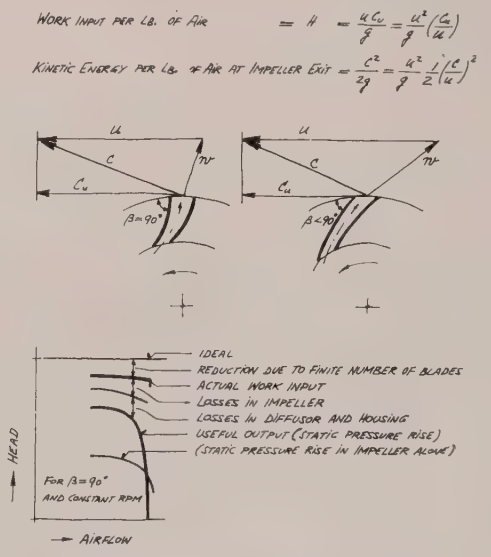

WORK INPUT PER LB. OF AIR $= H = \frac{u C_u}{g} = \frac{u^2}{g}\left(\frac{C_u}{u}\right)$
KINETIC ENERGY PER LB. OF AIR AT IMPELLER EXIT $= \frac{C^2}{2g} = \frac{u^2}{g}\frac{1}{2}\left(\frac{C}{u}\right)^2$
u
C
C_u
$\beta = 90°$
$\beta < 90°$
IDEAL
REDUCTION DUE TO FINITE NUMBER OF BLADES
ACTUAL WORK INPUT
LOSSES IN IMPELLER
LOSSES IN DIFFUSOR AND HOUSING
USEFUL OUTPUT (STATIC PRESSURE RISE)
(STATIC PRESSURE RISE IN IMPELLER ALONE)
HEAD
FOR $\beta = 90°$ AND CONSTANT RPM
AIRFLOW

Fig. 10 — Velocity triangle at impeller exit

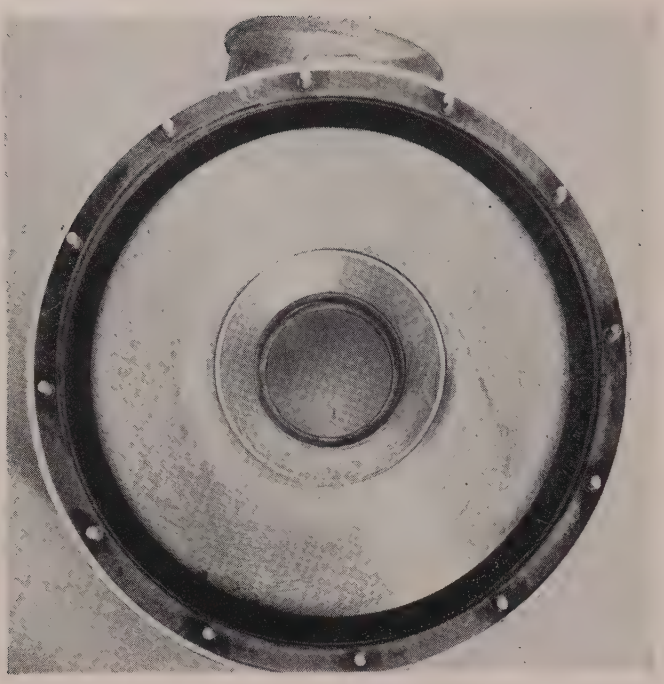


Fig. 11 — Flow pattern in vaneless diffuser

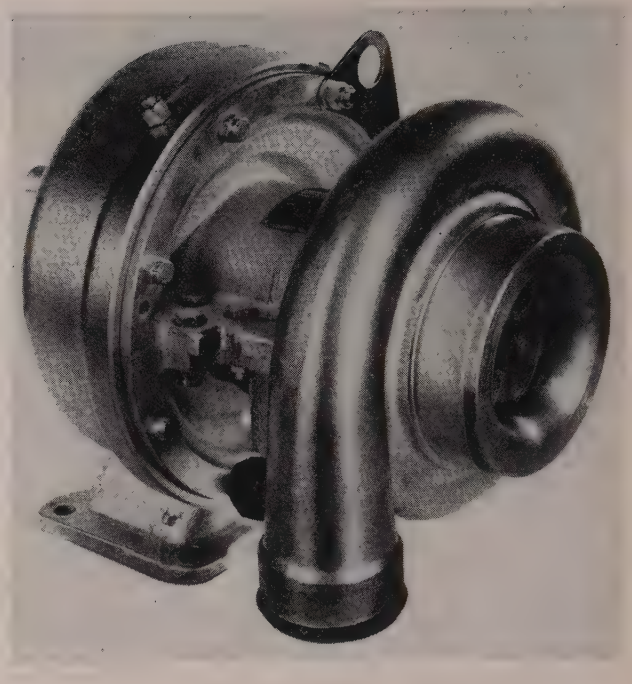


Fig. 12 — AiResearch turbocharger model T1407

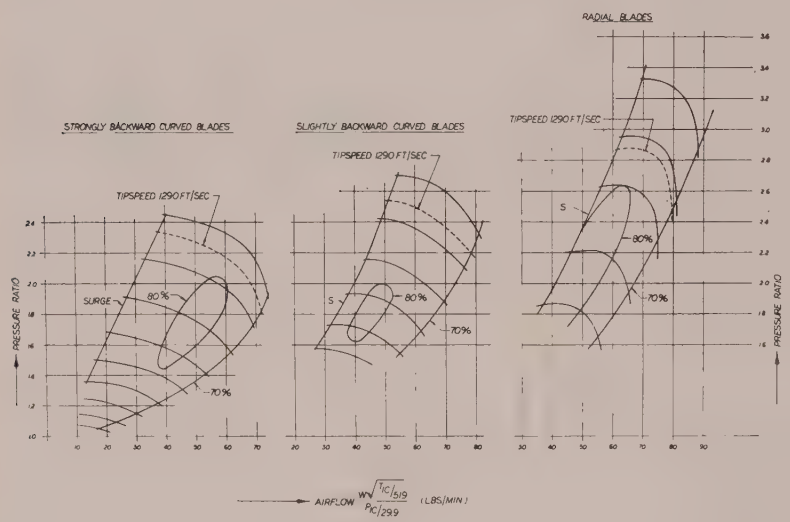


Fig. 13 — Influence of blade curvature on compressor map

useful output, of a 90-deg compressor running at constant speed, as a function of airflow.

Fig. 11 shows one side plate of a vaneless diffusor. The picture indicates clearly the spiraling motion of the air which has been made visible by injecting a lamp black — petroleum mixture into the compressor intake while running.

The air, after leaving the diffusor, has to be collected. Two types of housings have been widely used; volute type or torus type. Fig. 12 shows a volute housing.

Fig. 13 shows the performance maps of three different compressors. They are all of similar peak efficiency, have impellers of the same diameter, however, with different degrees of blade backward curvature, and have also approximately the same airflow at the same pressure ratio (similar gulp capacity). The operating ranges are bound at the left by the surge line and at the right, for comparative purposes, by a 70% efficiency line. The dotted speed lines correspond to the same impeller tip speed. The figure illustrates drastically the statements made above about range and pressure ratio. In fairness to the 90-deg compressor map shown, it should be said, however, that this compressor has been developed for a high compressor peak efficiency and that the range could be improved by sacrificing peak efficiency.

Referring back to Fig. 10, the work input equation shown there is very simple. It is only a conse-

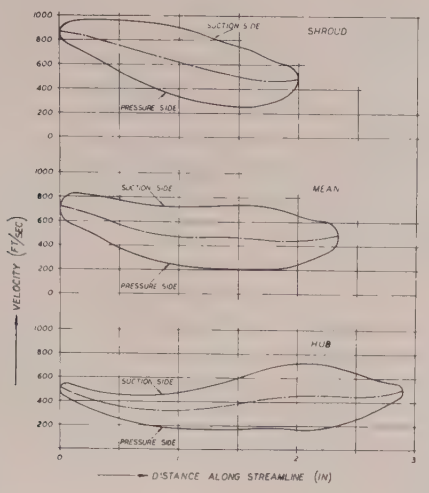


Fig. 14 — Velocity distribution along impeller blade

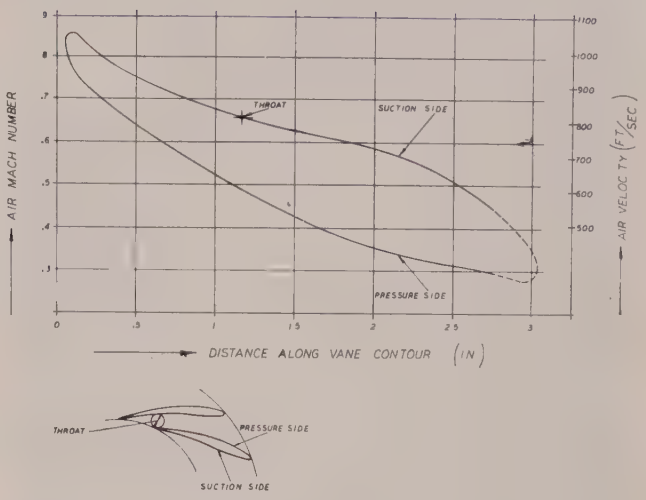


Fig. 15 — Velocity distribution along diffuser vane

quence of basic fluid mechanics and has nothing to do with aerodynamics. It does not give any information on how to achieve high efficiency and it does not help to determine optimum shapes of blades, meridional contours, blade numbers, and the like. Average performance turbomachines can be built with nothing more than just good basic mechanical engineering sense. To build high-performance machines of high gulp capacity and of fairly small absolute dimensions, however, a lot of aerodynamical work of an analytical and experimental nature is required. Even with today's aerodynamical know-how and electronic high-speed computing machines, there is a lot of "feel" left which makes an aerodynamicist a half-artist. Fig. 14, for instance shows calculated local velocity distributions within the impeller for its design point of rpm and airflow. The velocities are shown for three streamlines — shroud, mean, and hub, and for both sides of the blades — pressure and suction side. It is worthwhile to note the high velocities on the suction side along the shroud streamline which come close to the velocity of sound. Fortunately, today's computing machines permit quick calculations of velocity distributions of different geometrical shapes and for off-design points of operation from which a design can be chosen which appears most promising for manufacturing and testing of actual parts. Fig. 15 shows a similar velocity distribution for the suction and pressure side of a diffusor vane operating near the design point.

One of the consequences of the high tip speeds of turbomachines are the high stresses due to centrifugal forces as illustrated by Figs. 16 and 17. The figures show lines of constant radial and tangential stresses in the disc and hub of an aluminum impeller running at a rim speed which is well above the continuous duty operating speed. For this particular disc design, local yielding of the material takes place in the shaded area of Fig. 17. It is clear that the high tip speeds require a lot of develop-

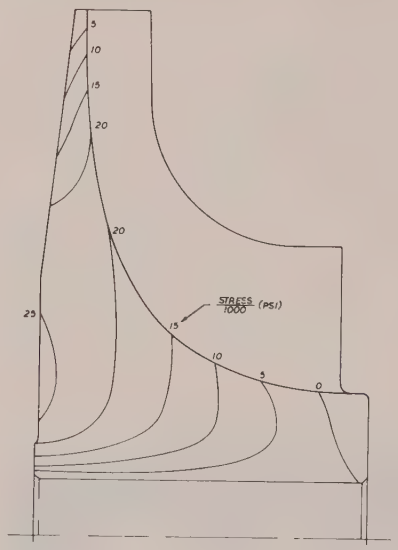


Fig. 16 — Radial stresses in impeller at overspeed

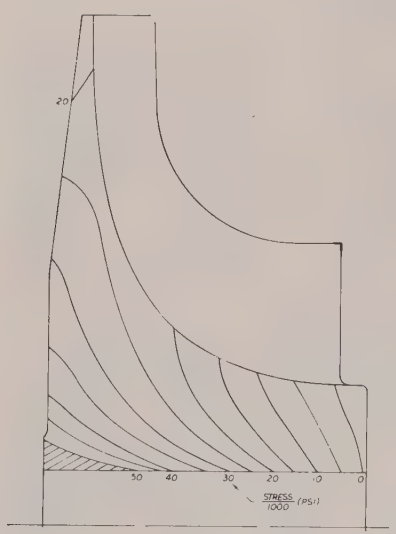


Fig. 17 — Tangential stress in impeller at overspeed

ment work and constant vigilance in order to produce sound impellers.

With the advent of very high pressure ratio turbocharging under continuous duty conditions, the impeller metal temperatures will increase too and the high temperature properties of impeller materials have to be considered quite carefully. Fig. 18 shows, for instance, two calculated life-to-rupture versus speed curves for an impeller based on long-time creep-to-rupture data of a particular aluminum alloy and foundry process. The two curves are for two different designs and are calculated taking into account the heat transfer between impeller and air which is heated while flowing through the impeller and undergoing its compression.

The stresses in the blades are also high. They are less subject to calculation because of the complicated blade shapes. Stress coat methods are used. In general, blades are more likely to fail by fatigue due to vibrations. Fig. 19 shows nodal lines of blade and disc vibrations as produced on a shaker table. Fig. 20 gives the whole spectrum of natural frequencies of the blades and the disc. An impeller

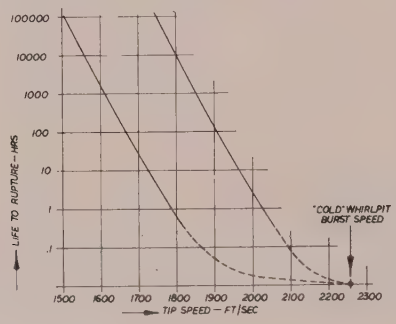


Fig. 18 — Typical life-speed relationship for aluminum impeller

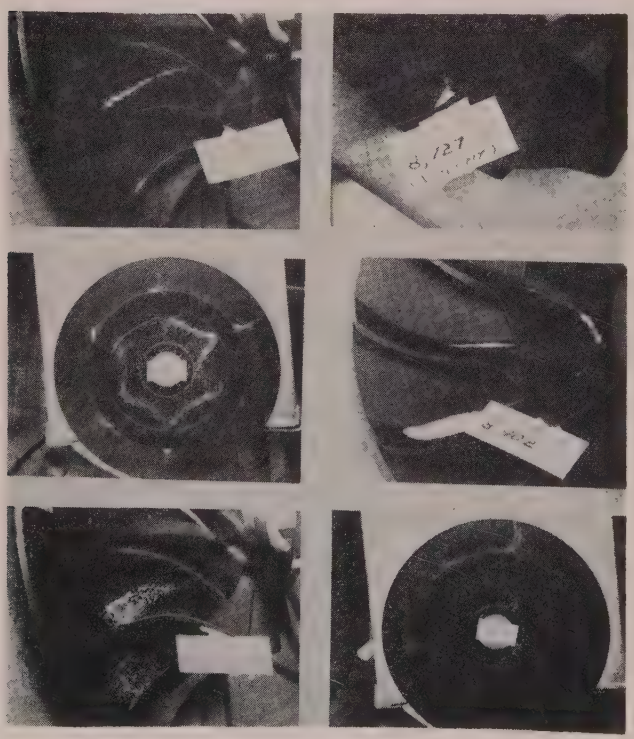


Fig. 19 — Nodal lines of blade and disc vibrations

with its blades represents a rather complicated part with a large number of different basic modes of vibration and coupled modes of vibration. The resemblance with a piano keyboard is striking. If one has tough luck one may always get resonance. Fortunately however, not all of these modes are very critical. Fig. 21 is a photo of an impeller with slightly backward curved blades with strain gages mounted in some critical areas for a test to record blade vibration amplitudes under running conditions.

Turbines and Turbocharger Size

The majority of turbochargers produced in this country have turbines of the radial or mixed flow type (type II and III of Fig. 9). Their aerodynamical and wheel stress problems are quite similar to the ones on the compressor side. This is no surprise since both types of wheels are so strikingly similar. As a matter of fact, the first radial gas turbine tests were done in Europe by using a radial compressor running in reverse with reversed flow direction, which gave excellent results.

The early radial gas turbines were truly radial, that is, the majority of the gas passage, as seen in a meridional plane, was in the radial direction with only a short axial portion at the exit called exducer. A turbine of this kind has a relatively large wheel diameter for a certain problem statement of gas flow, head, and discharge kinetic energy. In other words, it has a low gulp capacity. Since small size, weight, and moment of inertia are, in general, desirable, there has been a trend to wheels of higher gulp capacity with the result that today's turbines do not look like real radial turbines any more. The axial portion of the flow passage within the wheel has become predominant.

As indicated above, size considerations play an important role in the design of the turbine. This is because the turbine has to handle a volumetric flow rate of gas which can reach easily three times the volumetric flow rate of the compressor for a turbocharger used on a 4-stroke diesel engine. This is a consequence of the high exhaust temperature.

So far we have been mainly concerned with compressor density ratio, compressor efficiency, overall turbocharger efficiency, and impeller tip speed. It is now the time to look into what determines the size of a turbocharger. How small can a turbocharger be built for a certain given problem statement of air-flow and head? In order to arrive at the answer one has to consider both compressor and turbine.

As already illustrated in Fig. 14 the relative velocity of the air on the suction side of the impeller blades along the shroud contour close to the inlet, reaches very high values. Now, for a given problem statement of flow and head, and for a selected type of impeller blading characterized by its adiabatic head coefficient

$$q_{ad} = \eta_c \frac{H}{U^2/g} = \eta_c \left(\frac{C_u}{U} \right)$$

one can basically select any impeller diameter. (See also Fig. 10.) Then, there is one particular inducer diameter for which the relative approach velocity of the air at the inducer tip reaches a minimum. Obviously, the larger we choose the im-

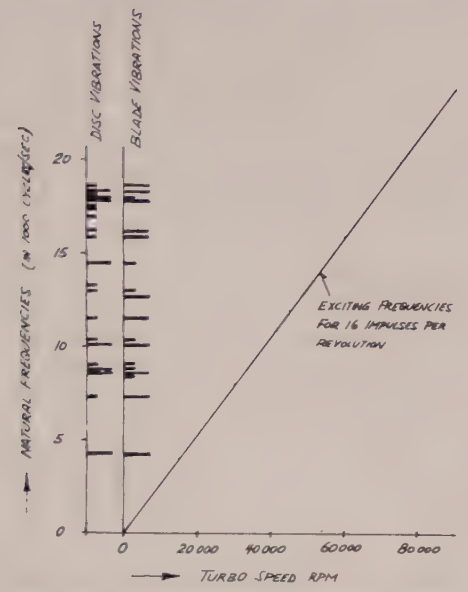


Fig. 20 — Spectrum of natural impeller vibration frequencies

Fig. 21 — Impeller with mounted strain gages

peller diameter the smaller the relative approach velocity and, therefore, conversely, the relative approach velocity determines the minimum impeller diameter. In order to have good performance also at off-design flow conditions (broad range), the approach velocity should not come too close to the velocity of sound. Introducing a Mach number M_c = relative approach velocity divided by inlet acoustic velocity, one finds for the first approximation of the minimum impeller diameter D_c:

$$D_c^2 \approx \frac{H_{ad_c}}{M_c^3} \qquad (2)$$

This result is independent of any turbine considerations. Now one can combine such a compressor of minimum diameter with a turbine. Again, one can investigate the importance of its exducer diameter D_{2T}. Assuming swirl-free gas flow at the turbine discharge one finds that with decreasing exducer diameter the kinetic energy in the exhaust gas at the exducer discharge increases continuously as shown in Fig. 22.

The turbine discharge loss is plotted on the abscissa in per cent of the compressor adiabatic head. This reference head has been selected to show the importance of the discharge loss as drastically as possible. The whole diagram is not universal, it has been calculated for a particular turbocharging problem. Again, the Mach number M_T of the relative velocity at the tip of the exducer is of great significance from the point of view of performance, and for the design of the blades and the optimum meridional shape of the wheel gas passage. Fig. 22 shows how the Mach number M_T first decreases and then increases for a continuously decreasing exducer diameter. The sketches at the bottom indicate turbine wheels corresponding to three different exducer diameters.

Apart from the two extreme regions at the end of the abscissa, the turbine Mach number is smaller than the compressor Mach number. This does not,

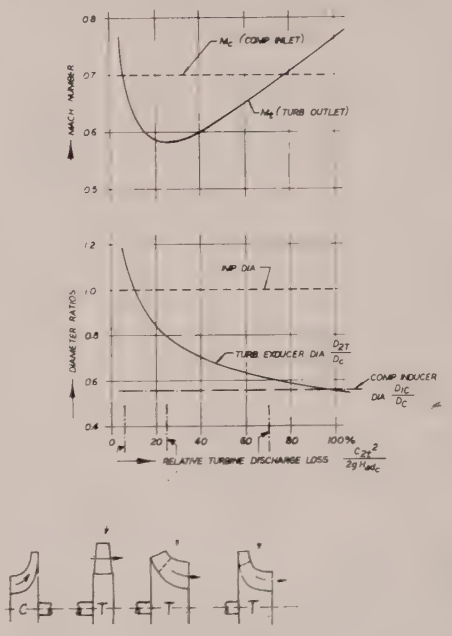


Fig. 22 — Different turbine wheel configurations matching minimum size impeller

however, imply that the aerodynamic problems in the turbine wheel are easier than in the compressor impeller. Since the turbine wheel is heavier than the aluminum impeller there will always be the tendency to make the turbine wheel as light as possible requiring aerodynamical compromises.

The quantity on the abscissa should actually have been called relative discharge energy and not relative discharge loss. A discharge diffusor after the turbine exit can recover a sizeable portion of this energy if designed properly and if the energy distribution at the exducer discharge is fairly uniform.

As the small sketches indicate, there are quite a number of design possibilities for the turbine (and

discharge diffusor), a field for exploration and evaluation of compromises between steady-state performance, weight, acceleration characteristics, production costs, and the like.

Fig. 23 gives a typical radial turbine performance map. The top portion shows the relationship between gas flow and pressure ratio for constant reduced speeds in between the two extremes of speed: runaway and stalled. The heavy line corresponds to the so-called "steam ellipse" which for the particular kind of selected reduced flow rate becomes a hyperbola. At the bottom are the efficiencies for different reduced speeds.

For a turbocharger with its compressor and turbine matched for high overall efficiency under steady-state conditions the operating points in the turbine flow versus pressure ratio map fall into a very narrow band along the "steam ellipse" line and, frequently, the influence of speed upon gulp characteristics can then be neglected in a first approximation.

The turbine map is used together with the compressor map to answer any questions regarding steady-state and off-design point turbocharger performance. It is also used in studying the acceleration behavior of the whole system.

When matching a turbine to a compressor the familiar presentation of turbine efficiency versus relative tip speed (tip speed divided by the head velocity $C_o = \sqrt{2gH_{adT}}$) as shown on Fig. 24 is very helpful. In order to keep the moment of inertia down, relative tip speeds slightly below the value for highest efficiency are usually selected. During fast acceleration the relative tip speed becomes real small because the turbine head is larger than for steady-state operation and the turbine efficiency is low as can be seen from the figure. Here is another area for possible progress.

Just a few words about the turbine nozzle. It is the counterpart of a vaned diffusor in a radial compressor. Its job is to create a swirling motion of the gas before entering the turbine wheel. In general, the gas increases its velocity while passing through the nozzle vanes. Because of weight reasons, and also because subdivided exhaust manifolds for best pulse utilization require small cross-sections for the turbine inlet port or ports, the inlet velocities to the turbine housing become very high. Sometimes their remains only a relatively small amount of flow acceleration for the nozzle passage. In the case of nozzles for radial turbines, this fact becomes even more pronounced and there is little work left for the nozzle vanes. It appears only natural in such a case to reduce the number of nozzle vanes or to dispose of the vanes entirely. In this last case, the turbine housing has to be shaped properly so it can take over the nozzle's function. Such a housing then becomes actually a one-vane nozzle.

Rotor Configurations

Of course, the best performing compressors and turbines are of no use if they are not properly integrated into a sound overall mechanical design concept of the whole unit. The basic arrangement is governed by the rotor configuration and the bearing location. In Fig. 25 today's most commonly found solutions are sketched.

The units with the straddle-supported rotor had their origin in Europe. Their original design stems largely from industrial compressor and steam turbine practice. Most of the early turbochargers fall into this category. This configuration is still very widely used abroad for large output engines. The bearings are far away from the impeller and turbine, leaving ample axial space for perfect seals for oil, air, and gas. The stubshaft diameters can be made small because they do not influence the critical speed, with the consequence of low bearing velocities and low bearing losses. All four ports for inlet and outlet of air and gas are in a radial direction. The structural connection between the two bearings has to go through or around these gas and air passages which is not easy on the hot turbine side. This difficulty does not, of course, exist if the turbine housing is water-cooled, which is the case in most of these designs. Water-cooling of the hot stationary parts makes a separate lube oil system simpler because of elimination of an oil cooler. Furthermore, water-cooling eliminates the difficulties of differential thermal expansion between compressor and turbine housings. The radial air and gas ports represent a handicap. In order to accommodate different engine installations, the four housing portions containing the four ports must have provisions for rotation. These requirements, together with the water-cooling, results in a heavy unit.

The single overhung rotor arrangement is used in this country for larger engines. This design can eliminate the disc friction, thereby improving the overall efficiency by a few points. Since the hot and cold portions of both the rotor and stator are built very closely together, design solutions had to be developed to take care of the associated problems of heat transfer, differential expansion, and distortion. The rotating assembly is characterized by its large overhang moment which required care-

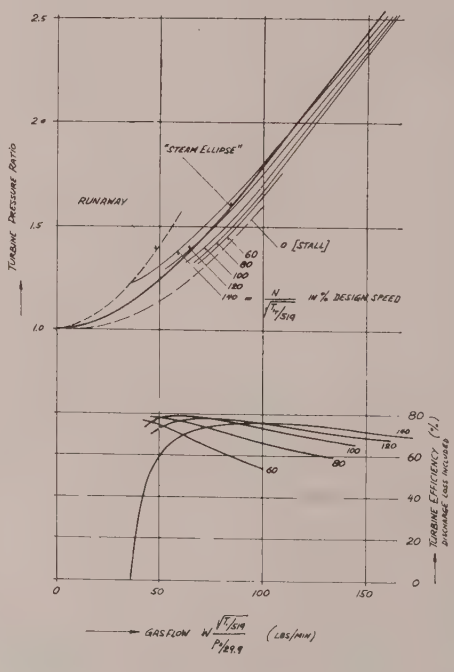


Fig. 23 — Typical performance map of radial turbine

ful investigation of the dynamical behavior of the rotor and bearing combination. The large diameter seal between the rotor blading of the compressor and the turbine controls the amount of leakage flow which, apart from being basically a small loss has, however, the advantage of contributing a cooling effect to the turbine rotor.

The next group of turbochargers with both the impeller and turbine wheel mounted overhung has found very widespread application here and abroad for smaller engines. This type of design imposes a difficult seal problem because of space limitations in order to keep the overhung moments small. Because the air inlet and the exhaust outlet are in an axial direction, only two housings have to have provisions for rotation to accommodate all kinds of installations. There is no necessity to carry a structural bearing support across any gas or air passage. These and the fact that both bearings are located within a common housing make this design really suitable for lightweight and low production cost. There is axial space available to provide for a suitable connection between "cold" compressor housing and hot turbine housing, permitting thermal differential expansion and sufficient protection of the compressor and the bearings from the turbine heat.

In mentioning turbine heat, Fig. 26 shows recorded bearing temperatures under a so-called soak-back test of a turbine with an overhung rotor without water cooling. This is a test where after fully stabilized high-temperature operating conditions, the gas supply to the turbine and the lube oil supply have been shut off simultaneously. The stored heat in the turbine wheel flows back into the shaft and into the bearings. It is interesting to note how two different designs of the center housing influence the temperature rise on the compressor bearing.

Needless to say that no matter what bearing design is used, hot turbo-machines do not appreciate this kind of treatment.

The last Fig. 27 shows a photograph of a cutaway model of one of the AiResearch Industrial Division's T14 high-pressure series turbochargers.

The information presented in this report gives a partial cross-section of the basic design principles of turbochargers. Since the main purpose of a turbocharger is to relieve the engine proper from the job of handling large volume flows, emphasis has been put on the flow handling characteristics of the turbocharger. No attempt has been made to describe activities in relation with the next very important subjects of any high-speed turbomachine, bearings, and associated dynamical behavior of the rotor in its bearings, lubrication and seals, and influence of production methods upon design.

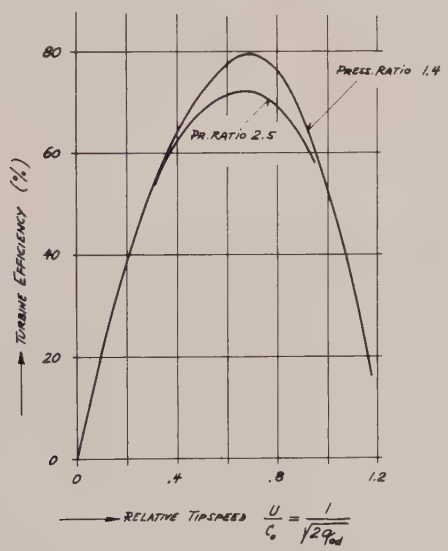


Fig. 24 — Turbine efficiency — speed relation

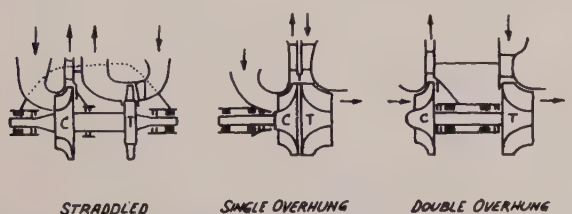


Fig. 25 — Most common rotor configurations

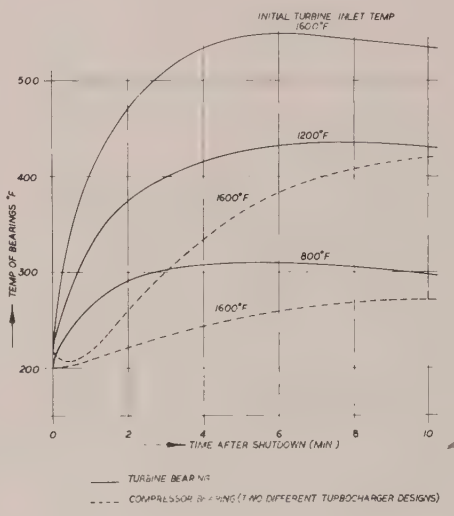


Fig. 26 — Bearing soak temperature

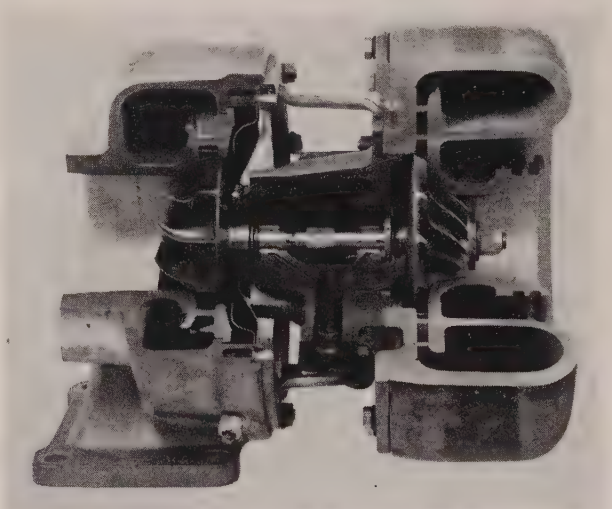


Fig. 27 — AiResearch T14 high-pressure series turbocharger

Relaxation of Residual Stresses due to Fatigue Loading

JoDean Morrow, University of Illinois
A. S. Ross, Boeing Airplane Co.
G. M. Sinclair, University of Illinois

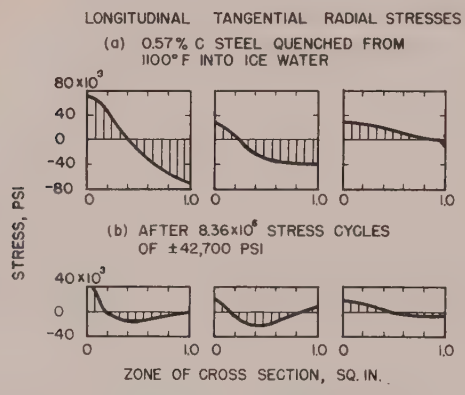


Fig. 1 — Change of residual stresses due to fatigue loading

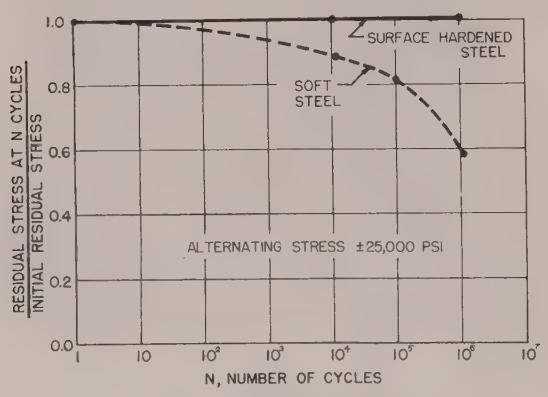


Fig. 2 — Cycle-dependent change of residual stresses in soft and hard steel

TO MANUFACTURE parts, metal is heated, deformed, abraded, hacked, or in some way forced to assume a useful shape. As a result, internal misfits exist in the part. Some regions simply do not fit into the space available. Internal stresses arise.

These residual stresses are usually greatest at the surface where the largest plastic deformation occurs during fabrication. Below the surface, complex triaxial states of stress may exist.

There are other complications; for example, the surface hardness and roughness of the part may be altered. Such parts often experience repeated loads in service which cause fatigue failures to initiate at the surface where the residual stresses are greatest.

Residual stresses have the effect of shifting the range of fatigue stressing much the same as mean stresses do in an axial fatigue test. Thus, compressive residual stresses are considered beneficial. Tensile residual stresses are detrimental.

The designer must insure satisfactory fatigue performance in the presence of still another complicating factor. Residual stresses in a member do not necessarily remain unchanged during its life. Possible sources of relief and redistribution of residual stresses in metals have been outlined by

MACRORESIDUAL stresses are usually induced in parts during fabrication. These stresses may change if the part is subjected to fatigue loading in service. To make an estimate of the effect of residual stresses on the life of the part, one must know the magnitude and importance of this change.

Axial fatigue results can be used to determine the mechanism by which residual stresses at the surface of a member relax due to fatigue loading.

Such an approach leads to the following conclusions:

1. Little relaxation of residual stresses should be expected for alternating stresses near the fatigue limit except for "soft" materials.

2. Because the fatigue limit of "soft" metals is, in general, relatively insensitive to mean stress, partial relaxation of residual stress will affect the fatigue limit only slightly.

3. For the purpose of estimating the fatigue limit of a member, the initial residual stress at the surface should be considered to remain unchanged during fatigue loading. Treating this stress as part of the service stresses, proceed as in any other fatigue problem in which mean stresses are present.*

Richards,[1] and one source is the fatigue loading itself. Yielding will occur at points where the sum of the residual stress and the applied stress exceeds the flow stress of the material. The material at these points will then "fit" better. Residual stresses change. The change will be largest early in the life of the part. Later, the change for a single cycle may be so small as to be imperceptible. However, millions of such cycles may cause a large cumulative effect.

It appears that, before the significance of residual stresses in fatigue can be appreciated, it is first necessary to know the influence of fatigue on residual stresses.

Object and Scope

Experimental investigations pertaining to the change of macroresidual stresses due to fatigue loading are summarized and interpreted. The mechanism by which residual stress is believed to change is formulated, and an attempt is made to answer these questions:

1. What determines the amount of change?

2. How should the designer account for this change?

Some Experimental Evidence

Residual stresses may be decreased, increased, or left unchanged by fatigue loading. For example, Fig. 1 shows the decrease of triaxial residual stresses by Bühler and Buchholtz.[2] Residual stresses were measured before and after fatigue testing (8.36×10^6 cycles), but no attempt was made to determine the cyclic rate of change.

Change in residual stress is largest for the first few cycles. In fact, Wallace and Frankel[3] found that the residual stress in a structural steel member was removed by a single cycle of fatigue loading.

Moore[4] measured the percentage change of residual stresses as a function of the number of fatigue cycles. These results are shown in Fig. 2. Here the fraction of the initial residual stress remaining after a given number of cycles is plotted versus the log of the number of cycles. Notice that no relaxation occurred in the surface hardened steel. In the soft structural steel, there was a progressive decrease in the residual stress as the number of cycles increased.

The progressive nature of the relaxation of residual stresses due to fatigue loading received considerable attention by a Japanese investigator. Taira[5] induced surface compressive residual stresses of about 25,000 psi by prestraining unnotched soft steel specimens. These were subjected to repeated bending loads. The residual stresses were determined periodically at the surface by X-ray. Taira's results are shown in Fig. 3. Notice that the rate of relaxation depends greatly upon the amplitude of alternating stress. After 10^4 cycles the plot of the fraction of the original residual stress remaining versus log of the number of cycles is practically linear.

Elsesser and Corten[6] report "... an increase in the

---

* Paper presented at meeting of Division 4 — Residual Stresses, SAE Iron and Steel Technical Committee, Detroit, Jan. 15, 1959.

[1] "Relief and Redistribution of Residual Stresses in Metals," by D. G. Richards. Pp. 129–191, "Residual Stress Measurements." Pub. by American Society for Metals, Cleveland, 1951.

[2] "Effect of Residual Stresses upon Bending Fatigue Strength," by H. Bühler and H. Buchholtz. *Stahl und Eisen*, Vol. 53, 1933. Brutcher Translations No. 53, pp. 1330–1332.

[3] "Relief of Residual Stress by Single Fatigue Cycle," by W. P. Wallace and J. P. Frankel. *Welding Journal*, Vol. 28, November, 1949, p. 565s.

[4] "Study of Residual Stresses and Size Effect and Study of Effect of Repeated Stresses on Residual Stresses due to Shot Peening of Two Steels," by H. F. Moore. Proceedings, Experimental Stress Analysis, Vol. 2, No. 1, 1944, pp. 170–177.

[5] "X-Ray Studies of Change of Residual Stress due to Stress Repetition," by S. Taira. Japan Society of Mechanical Engineers Proceedings, Vol. 1, 1948, p. 107; and Vol. 16, 1950, p. 2 (in Japanese).

[6] "Influence of Repeated Loads on Residual Stresses in Inelastically Deformed Beams," by T. M. Elsesser and H. T. Corten. Paper 54-F-13 presented at ASME Fall Meeting, Milwaukee, September, 1954.

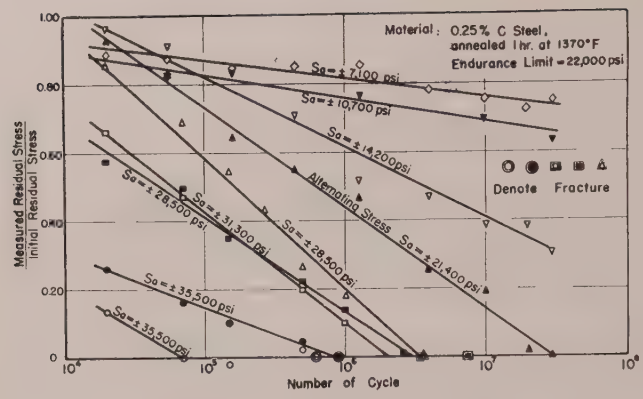

Material: 0.25% C Steel, annealed 1 hr. at 1370°F
Endurance Limit = 22,000 psi
Sa = ± 7,700 psi
Sa = ± 10,700 psi
Sa = ± 14,200 psi
Sa = ± 21,400 psi
Sa = ± 28,500 psi
Sa = ± 31,300 psi
Sa = ± 35,500 psi
Sa = ± 35,500 psi
Alternating Stress
Denote Fracture
Measured Residual Stress / Initial Residual Stress
Number of Cycle

Fig. 3 — Decrease of residual stress due to repeated loading

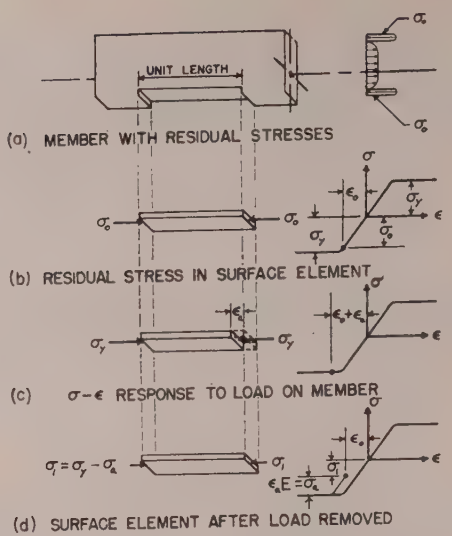

UNIT LENGTH
(a) MEMBER WITH RESIDUAL STRESSES
(b) RESIDUAL STRESS IN SURFACE ELEMENT
(c) σ–ε RESPONSE TO LOAD ON MEMBER
(d) SURFACE ELEMENT AFTER LOAD REMOVED

Fig. 4 — Illustration of residual stress change

Table 1 — Test Program

Material	R_c Scale	No. of Specimens in Set	Initial Mean Stress, psi	Approximate Fatigue Limit, psi
SAE 4340	29	15	0	57,000
SAE 4340	29	13	60,000	45,000
SAE 4340	29	13	100,000	36,000
SAE 4340	29	9	– 60,000	66,000
SAE 4340	41	6	0	60,000
SAE 4340	41	10	60,000	42,000
SAE 4340	41	12	140,000	30,000
SAE 4340	50	18	0	80,000
SAE 4340	50	14	60,000	60,000
SAE 4340	50	21	120,000	20,000
SAE 4340	50	7	180,000	15,000
Alloy A-286	25	7	0	58,000
Alloy A-286	25	11	30,000	55,000
Alloy A-286	25	9	60,000	46,000
Alloy A-286	25	8	80,000	41,000
Alloy A-286	25	9	– 60,000	65,000

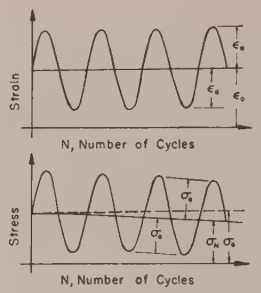


Fig. 5 — Schematic test conditions

residual stresses in the outermost fibers of the beams with increasing number of load repetitions." These tests were performed on SAE 1035, 1045, and 4340 steels. Initially, stress-free beams were loaded from zero to maximum. Rosenthal and Sines[7] also found an increase in the residual stress at the root of a notch in aluminum specimens subjected to zero to maximum loading.

For hard steel (R_c 59), Tarasov, Hyler, and Letner[8] observed, "Cyclic stressing did not cause any measurable reduction in the magnitude of the residual grinding stresses." This comment was based on cyclic tests with applied stresses near the fatigue limit.

The experimental evidence indicates that these factors determine the change of residual stress during a fatigue test:

1. Composition and hardness of the material.
2. Magnitude and sense of the initial stresses present.
3. Magnitude and direction of the applied fatigue stress (whether reversed or unidirectional).
4. Number of cycles of fatigue loading.

There have been no comprehensive experiments to determine directly the relative importance of these factors. Many experimental difficulties make it impractical to carry out such a program. For example, methods of inducing residual stresses such as shotpeening or quenching may alter the surface material drastically. Techniques for determining residual stresses are tedious and time consuming, making extensive testing prohibitive.

In the next section an experimental approach is described which overcomes some of these difficulties.

Experimental Approach

Fig. 4(a) shows a member with longitudinal residual stresses. To simplify the problem, consideration will be given only to the surface element shown in Fig. 4(b), which is held in place by the surface residual stress, σ_o. This means that the surface element would be longer by the amount, ϵ_o, if it were removed from the surrounding material.

Presume the material has a stress-strain curve as shown in the right portion of Fig. 4(b). The amount of stress and strain in the element appears as a point on the stress-strain diagram.

Apply an axial load or bending moment to the

[7] "Effect of Residual Stress on Fatigue Strength of Notched Specimens," by D. Rosenthal and G. Sines. ASTM Proceedings, Vol. 51, 1951, p. 593.
[8] "Effect of Grinding Conditions and Resultant Residual Stresses on Fatigue Strength of Hardened Steel," by L. P. Tarasov, W. S. Hyler, and H. R. Letner. ASTM Proceedings, Vol. 57, 1957, p. 601.

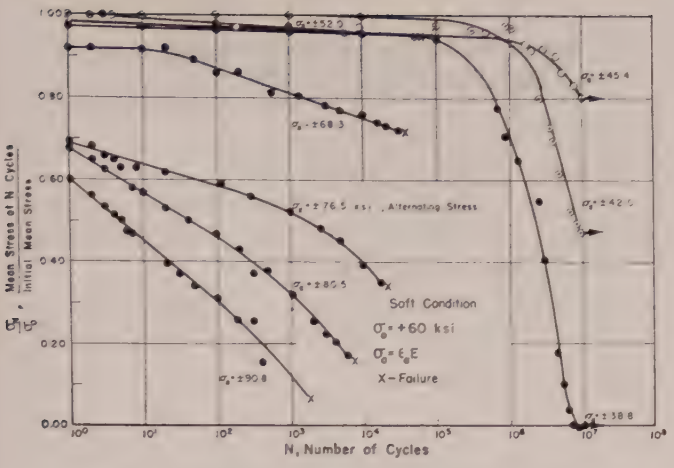


Fig. 6 — Relaxation of mean stress of SAE 4340 steel

member. Choose the direction of loading such that the element shown in Fig. 4(c) is subjected to the compressive strain, ϵ_a. Provided $\epsilon_0 + \epsilon_a$ is large enough, yielding will occur in the surface element. The stress on the surface element will then be σ_y, the flow stress of the material.

Unload the member. If only a small region of the total cross-section experiences yielding, the portion which remains elastic will restore the strain in the surface element near the original value, ϵ_0. The stress will decrease to the value σ_1 as shown in Fig. 4(d).

Repeated loading of the member would cause additional relaxation of the surface residual stresses.

Why not consider the surface element of Fig. 4 to be analogous to a uniaxially loaded fatigue specimen? Initial stresses and strains in such an axially loaded filament are comparable to the initial stresses and strain present in the surface elements of members with residual stresses. The amplitude of alternating strain imposed on the axial fatigue specimen is analogous to the repeated strains caused in the surface element of the member due to fatigue loading. The change of stress from the original value due to the fatigue loading is similar to the change of residual stress in the member.

Justification for using this simple approach to study the change of residual stresses in members subjected to fatigue loading are as follows:

1. Fatigue failure usually nucleates at the surface. Thus, a knowledge of what happens at the surface is sufficient to determine the fatigue behavior of the part.

2. Service stresses are normally not large enough to cause excessive yielding of the part as a whole. Yielding takes place at discreet regions where stresses are highest. Therefore, it is possible to use the simplifying assumption that surface elements experience repeated strain when the member is subjected to repeat loads.

3. Inducing stress in this manner does not alter the surface finish or the properties of the material being tested. Basic information is obtained for separating the effect of work hardening, surface finish, and residual stresses.

4. No complicated measurement of residual stress is needed, permitting large amounts of data to be obtained with relative ease. The separate effect of the factors which influence the relaxation of residual stress can be studied exhaustively.

Such an experimental program and the resulting analysis are reported in the next two sections.

Cycle-Dependent Stress Relaxation Tests

There follows a discussion of axial fatigue tests on SAE 4340[9] and the alloy A-286.[10] The mean strain and amplitude of alternating strain were held constant throughout the fatigue tests, as shown in Fig. 5. These conditions were intended to approximate the conditions of loading of the surface element of Fig. 4 discussed in the preceding section.

Under these conditions, any inelastic action accumulated would cause the mean stress to decrease, as also shown in Fig. 5. The relaxation or change of the mean stress is considered similar to the change of residual stresses in members subjected to fatigue loading. Since the stress depends on the number of cycles, this phenomenon has been called cycle-dependent stress relaxation.

Three levels of hardness were investigated for the SAE 4340 steel (R_c 30 — soft, 40 — medium, 50 — hard), and the hardness of the A-286 was C 25. Initial mean stresses ranging from 60,000 psi in compression to 180,000 psi in tension were investigated. Table 1 summarizes the experimental program and gives some of the pertinent data. For further detail the reader is referred to footnotes 9 and 10. Figs. 6–11 are samples of cycle-dependent stress relaxation data discussed below.

Discussion of Sample Relaxation Data — In Fig. 6 the fraction of the initial mean stress remaining at N cycles is plotted against the log of N. The material is SAE 4340 in the soft condition with an initial tensile mean stress of 60,000 psi. Each line represents the results of one specimen tested at a particular value of alternating stress.[11] The curves shown in this figure are simply to guide the eye; they do not represent theoretical or expected behavior. On the semi-log plot used, these curves are nearly linear below N equal to approximately 10^6 cycles.

For alternating stresses which cause fatigue fracture to occur, there is a decrease of mean stress during the first cycle which continues as the number of cycles increases. For large amplitudes of alternating stress, the first cycle decrease of mean stress and the slope of the relaxation line on the semi-log used are large.

For smaller amplitudes there is little relaxation of mean stress in less than one million cycles. After a million cycles there is sometimes a decrease of the mean stress even for specimens which endure 10^7 cycles.

The test conditions for the data displayed in Fig. 7 are identical to those in Fig. 6 except for the initial

[9] "Cycle-Dependent Stress Relaxation," by JoDean Morrow and G. M. Sinclair. Special Technical Publication 237, Symposium on Basic Mechanisms of Fatigue, ASTM, 1958. Also, T & AM Report No. 543, University of Illinois, August, 1957.

[10] "Cyclic Stress Behavior of A-286 Alloy for Conditions of Controlled Strain," by A. S. Ross and JoDean Morrow. T & AM Report No. 563, University of Illinois, September, 1958. (Presented at ASME Annual Meeting, 1959.)

[11] For convenience in this paper, the term alternating stress is used for the product of the alternating strain and elastic modulus, $\epsilon_a E$. The *actual* alternating stress is less than $\epsilon_a E$ early in the test but approaches this value after the first few cycles.

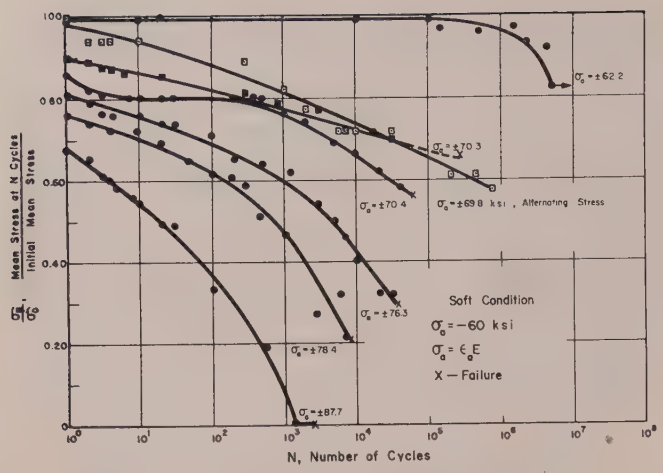

Mean Stress at N Cycles / Initial Mean Stress
N, Number of Cycles
Soft Condition
$\sigma_0 = -60$ ksi
$\sigma_a = \epsilon_a E$
X — Failure
$\sigma_a = \pm 69.8$ ksi, Alternating Stress

Fig. 7 — Relaxation of mean stress of SAE 4340 steel

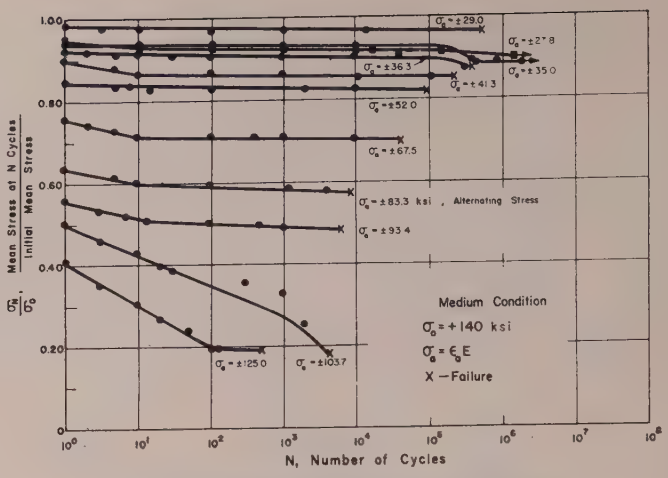

Mean Stress at N Cycles / Initial Mean Stress
N, Number of Cycles
Medium Condition
$\sigma_0 = +140$ ksi
$\sigma_a = \epsilon_a E$
X — Failure
$\sigma_a = \pm 83.3$ ksi, Alternating Stress

Fig. 9 — Relaxation of mean stress of SAE 4340 steel

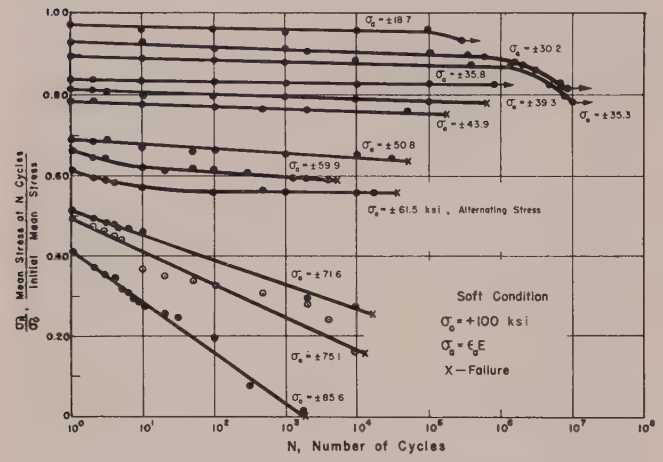

Mean Stress at N Cycles / Initial Mean Stress
N, Number of Cycles
Soft Condition
$\sigma_0 = +100$ ksi
$\sigma_a = \epsilon_a E$
X — Failure
$\sigma_a = \pm 61.5$ ksi, Alternating Stress

Fig. 8 — Relaxation of mean stress of SAE 4340 steel

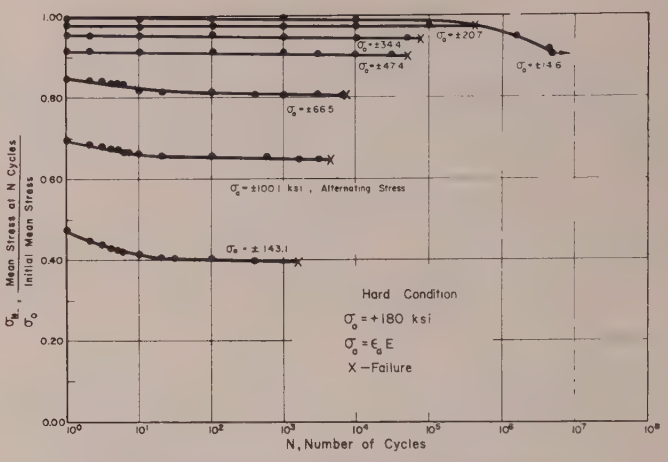

Mean Stress at N Cycles / Initial Mean Stress
N, Number of Cycles
Hard Condition
$\sigma_0 = +180$ ksi
$\sigma_a = \epsilon_a E$
X — Failure
$\sigma_a = \pm 100.1$ ksi, Alternating Stress

Fig. 10 — Relaxation of mean stress of SAE 4340 steel

mean stress which was in compression. Comparison of the two figures reveals that the behavior is nominally the same.

Figs. 8–10 give cycle-dependent relaxation data for the soft, medium, and hard conditions of SAE 4340. The initial mean stresses for these three figures are approximately 80% of the 0.2% offset yield strength for the three conditions of hardness. The influence of hardness is evident from these figures. For the soft condition, the relaxation occurs more readily than for the hard condition while for the medium condition, the behavior is intermediate. As in Figs. 6 and 7, very little relaxation of mean stress took place before 10^6 cycles for specimens which did not fail.

Note the similarities between Figs. 6–10 for cycle-dependent stress relaxation and Figs. 2 and 3 for change of residual stresses due to fatigue loading.

Alloy A-286 — A sample of the relaxation of mean stress obtained for alloy A-286 is given in Fig. 11. While these results agree qualitatively with the results for SAE 4340 steel, a quantitative comparison is a bit disappointing. The A-286 used in this investigation was softer than the soft SAE 4340. On this basis, one would expect the relaxation of mean stress to be larger in this material for the same amplitude of alternating stress. Yet, a comparison of similar tests from Fig. 11 for A-286 and Fig. 8 for the soft SAE 4340 reveals that the SAE 4340 relaxed more than A-286 for the same alternating stress. In fact, for A-286 there is virtually no relaxation of mean stress after the first ten cycles.

Little relaxation of the mean stress occurred below 10^6 cycles for A-286 specimens which did not fail. For both SAE 4340 and A-286 at least 75% of the initial mean stress was still present after a million cycles for those specimens which did not fail in fatigue.

Analysis of Cycle-Dependent Stress Relaxation

Cycle-dependent relaxation is a result of material yielding. On the first cycle, yielding may be quite large. On ensuing cycles, yielding is small. It is, therefore, convenient to break the problem of analysis into two parts — first cycle relaxation and relaxation for later cycles.

Mean Stress After One Cycle—In Fig. 12 the stress-strain response to the first cycle of alternating strain about a fixed mean strain is shown schematically. Assume the material to have a "flat top" stress-strain curve with a yield point, σ_y, and elastic unloading along a line parallel to the elastic slope E. The following expression for the mean stress at

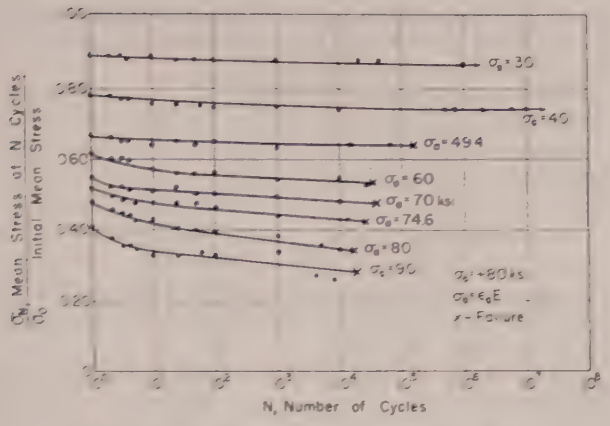


Fig. 11 — Relaxation of mean stress of A-286

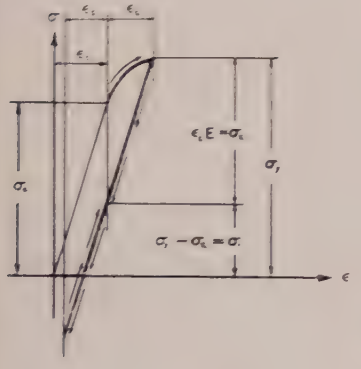


Fig. 12 — First cycle change of mean stress

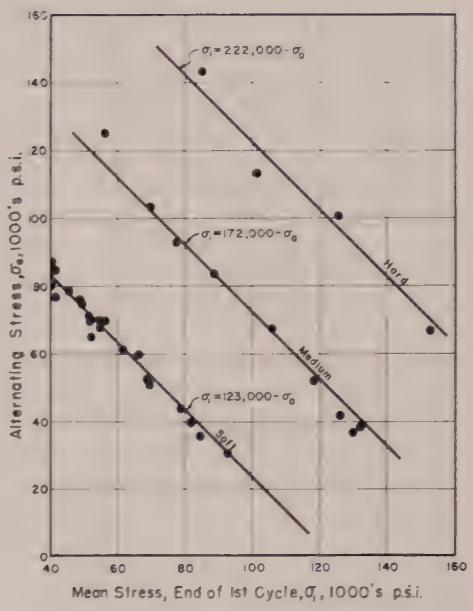


Fig. 13 — Mean stress after one cycle

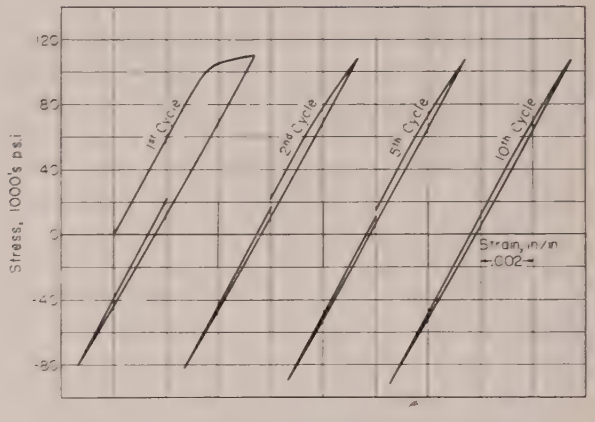


Fig. 14 — Cyclic stress-strain behavior of A-286

the end of the first cycle, σ_1, can be obtained from inspection of the figure:

$$\sigma_1 = \sigma_y - \epsilon_a E = \sigma_y - \sigma_a \qquad (1)$$

where:

ϵ_a = Amplitude of alternating strain

$\sigma_a = \epsilon_a E$

In Fig. 13 the mean stress at the end of the first cycle is plotted versus the alternating stress for all the SAE 4340 specimens in which large-scale yielding occurred on the first cycle. Three lines are placed through the data for the soft, medium, and hard specimens. The lines are of the form of Eq. 1. The following values of σ_y were used:

Soft	123,000 psi
Medium	172,000 psi
Hard	222,000 psi

These values are slightly less than the 0.2% offset yield strength for the three hardness levels given in Table 1.

Good correlation of the mean stress at the end of the first cycle on the basis of Eq. 1 is possible for SAE 4340, as Fig. 13 demonstrates.

For A-286 no simple equation for correlating the mean stress at the end of the first cycle was found. This material demonstrated stress-strain properties which were more complex than those in the simple model shown in Fig. 12, from which Eq. 1 was derived. Stress-strain loops for A-286 at 1, 2, 5, and 10 cycles are shown in Fig. 14. Notice that two of the assumptions which were made in the derivation of Eq. 1 are violated during the first cycle. The stress-strain curve is *not* "flat-top" and the material does *not* unload linearly along a line parallel to the elastic modulus line. Due to these factors, the mean stress at the end of the first cycle was found always to be above the value predicted from Eq. 1. The larger the strain amplitude and the initial mean stress, the larger the difference. An analysis of the first cycle relaxation taking these factors into account is given in Ross and Morrow.[10]

No appreciable change in the mean stress occurred during the first cycle for either the SAE 4340 or the A-286 if the sum of the alternating stress and the initial mean stress was below a value roughly equal to the 0.2% offset yield strength.

Relaxation of Mean Stress for Cycles Subsequent to First — After the first cycle there should be no large scale plastic yielding for the condition of controlled strain used in this investigation. However, there might be microinelastic action which could be observed only by accumulating these small

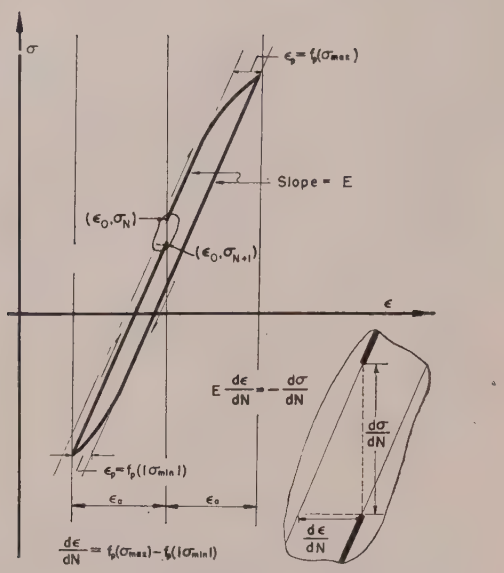


Fig. 15 — Assumed change of mean stress on N^{th} cycle

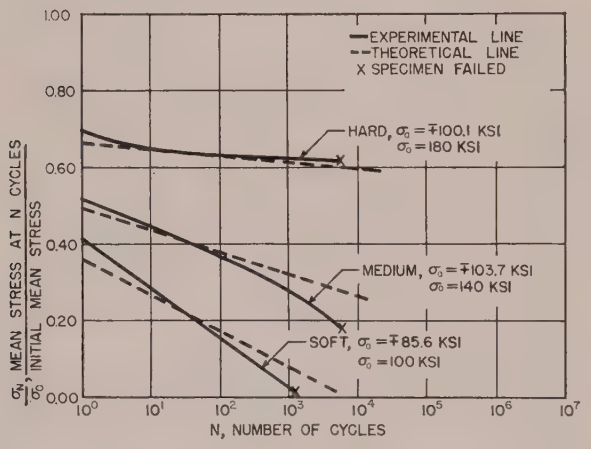

EXPERIMENTAL LINE
THEORETICAL LINE
X SPECIMEN FAILED
HARD, $\sigma_0 = \mp 100.1$ KSI
$\sigma_b = 180$ KSI
MEDIUM, $\sigma_0 = \mp 103.7$ KSI
$\sigma_b = 140$ KSI
SOFT, $\sigma_0 = \mp 85.6$ KSI
$\sigma_b = 100$ KSI
σ_N / σ_0, MEAN STRESS AT N CYCLES / INITIAL MEAN STRESS
N, NUMBER OF CYCLES

Fig. 16 — Comparison of computed and observed behavior of SAE 4340 steel

strains over a large number of repeated strain cycles. A helpful model for considering this type of microplacity appears in Fig. 15. Here the deviation from linearity is exaggerated for clarity.

For the *N*th cycle (Fig. 15), the material is presumed to be strained along some nonlinear path to the maximum strain. It is then unloaded elastically to zero stress. It is presumed to follow the same shape path when loaded in compression as was followed in tension. Returning the strain to the initial mean value completes one strain cycle. The material experiences the small plastic strain $d\epsilon/dN$ which causes the stress-strain loop *to fail to close* by the amount $d\sigma/dN$. From the enlarged view at the bottom right of the figure, $d\epsilon/dN$ can be related to $d\sigma/dN$ as follows:

$$E\frac{d\epsilon}{dN} = -\frac{d\sigma}{dN}$$

Further study of this assumed stress-strain loop reveals that $d\epsilon/dN$ is the difference between the plastic strain at the maximum stress and the plastic strain at the minimum stress.

Assume a function of stress, $f_p(\sigma) = \epsilon_p$, which determines the amount of plastic strain for a given stress level.[12] The following equation can be written for the mean stress at *N* cycles:

$$-\frac{1}{E}\frac{d\sigma_N}{dN} = f_p(\sigma_{\max}) - f_p(|\sigma_{\min}|) \qquad (3)$$

To solve Eq. 3 a function must be chosen which represents realistically the cycle stress-plastic strain behavior of the material.

Such a solution has been made in Morrow and Sinclair[9] where the resulting equations can be found. In the same reference it is shown that the following approximate equation can be obtained by making certain simplifying approximations:

$$\frac{\sigma_N}{\sigma_o} = \frac{\sigma_1}{\sigma_o} - \left(\frac{\sigma_a}{\sigma_y}\right)^b \log N \qquad (4)$$

where:

σ_N = Mean stress at *N* cycles
σ_o = Initial mean stress
b = Material constant dependent on the cyclic stress-strain properties

Eq. 4 is not applicable for values of σ_N/σ_o less than about 0.20, nor for *N* greater than about 10^6 cycles.

Fig. 16 has been prepared to compare Eq. 4 to experimental results. The bold lines are reproductions of representative experimental relaxation curves for SAE 4340 taken from Figs 8–10, and the dashed lines are theoretical curves obtained with the aid of Eqs. 1 and 4 with $b = 5.7$ and σ_y, defined as before. The agreement can be seen to be good. Space does not permit graphical comparison of more data. However, it was found that the experimental values of σ_N/σ_o were within about ± 0.10 of the value from Eq. 4 for all SAE 4340 data reported in Morrow and Sinclair,[9] except for $N > 10^6$ and for values of $\sigma_N/\sigma_o < 0.20$.

As mentioned before, the cycle-dependent relaxation behavior of A-286 was substantially different than that of SAE 4340. Most of the relaxation occurred in the first 10 cycles. Reasons for the difference in the behavior of SAE 4340 and A-286 can be found in Fig. 14. After a few cycles of fatigue loading, the cyclic stress-strain behavior of A-286 is such that the hysteresis loop tends to join in a continuous loop. (See the tenth cycle, Fig. 14.) Under these conditions the model used for relaxation (Fig. 15) is no longer applicable since $d\sigma/dN$ approaches 0. The hysteresis loop reaches a stable shape which does not permit further relaxation of mean stress under conditions of controlled strain.

Despite the absence of cycle-dependent relaxation in A-286, it was evident that a broad hysteresis loop was present because specimens tended to heat excessively during testing.

Further details concerning cycle-dependent stress relaxation are not appropriate to the purpose of this paper. The question is — how can this information be used to understand the change of residual stresses in parts which are subjected to

[12] This function is assumed to be the same in tension as in compression and to remain unchanged by repeated straining of the material.

fatigue loading? This question is taken up in the next section.

Change of Residual Stresses Due to Fatigue Loading

The surface element of a member containing residual stresses has been considered to be analogous to an axially loaded fatigue specimen. On the basis of this analogy, the factors which are considered to determine the amount of change of residual stresses due to fatigue loading will be discussed.

Magnitude and Sense of Initial Residual Stresses and Applied Stresses — As Dolan[13] suggested, residual stresses should be considered "... a portion of the imposed loading pattern that the material must resist." For all types of loading the surface element is subjected to an alternating stress about a mean stress.[14] The mean stress will always *tend* to relax toward zero when the member is subjected to fatigue loading. That is, yielding or cycle-dependent relaxation will cause the cyclic stress at the surface to approach conditions of completely reversed stressing (mean stress = zero).

Surface residual stress in a member, such as Fig. 4 shows, is analogous to mean stress if the member is subjected to completely reversed loading. For the case of one directional loading of the member the residual stress will be the maximum or minimum stress in the cycle depending on its sense and the direction of the applied load. If the member is subjected to a mean and an alternating load, the residual stress will not be the mean, nor will it be the maximum or minimum stress.

A few examples will be used to demonstrate why residual stresses can decrease, increase, or remain unchanged:

1. Completely reversed loading.

Under these conditions the residual stress is the mean stress. Should the sum of the residual stress and the alternating stress at the surface be large enough, the residual stress will relax toward zero. Thus, the residual stress decreases. Examples have been cited from the literature and from the experimental program reported here in which such a decrease was observed.

2. Zero initial residual stresses — one directional loading.

The stress-strain response of the surface element is shown in Fig. 17. While the mean stress relaxes toward zero, the residual stress at the surface is changed from zero to a compressive value. Should the mean stress reach zero eventually, the residual stress at the surface of the member would be equal to the amplitude of the alternating stress. A residual stress induced by cycling an initially stress free bar has been reported by Elsesser and Corten[6] and is discussed earlier in this paper.

It is also possible for an initial residual tensile stress to be changed to a residual compressive stress in the same manner as shown in Fig. 17. Or compressive stresses can be changed to tension.

3. High residual stress present, but loaded such that the mean stress is zero.

These conditions are shown in Fig. 18. There is a large residual stress present, and the range of applied stress is large. The mean stress is zero. The stress-strain behavior of the surface material is such that no change occurs in the residual stress. Strain peening[15,16] is based on this principle. It insures dependable performance of leaf springs in service by inducing residual stresses of opposite sense to the applied stresses.

Hardness — For the same material the residual stress will be changed by fatigue loading more easily in the soft state than in the hard state, other conditions being equal. However, hardness alone is not a sufficient index. More significant are the cyclic stress-strain characteristics of the material, which may be determined by tests similar to those reported above.

It is important to remember that the induction of residual stresses at the surface of a member may be accomplished by work hardening or thermal softening. Therefore, a process such as shotpeening would not only induce desirable compressive residual stresses, but would tend, also, to insure that the residual stresses do not change as easily because of the increase in the hardness of the surface material.

Number of Load Repetitions. — Most of the change in the residual stress will occur during the first few cycles. For sufficiently large ranges of

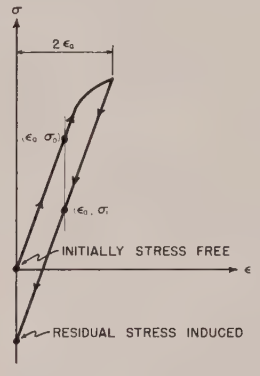


Fig. 17 — Residual stress induced into stress-free member by fatigue loading

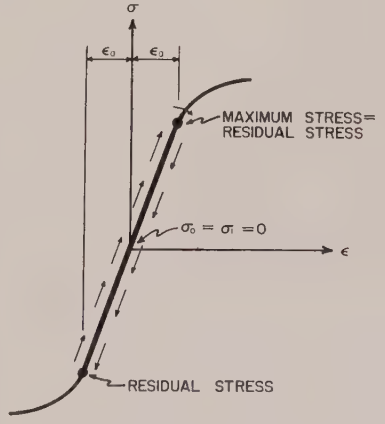


Fig. 18 — Residual stress does not change if mean stress is zero

[13] "Residual Stress, Strain Hardening and Fatigue," by T. J. Dolan. From "Internal Stresses and Fatigue in Metals." Pub. by Elsevier Publishing Co., Amsterdam, Holland.

[14] The amplitude of alternating load is considered to remain constant.

[15] "Effect of Shot Peening Variables and Residual Stresses on Fatigue Life of Leaf Spring Specimens," by R. L. Mattson and W. S. Coleman, Jr. SAE Transactions, Vol. 62, 1954, pp. 546–556.

[16] "Effect of Residual Stresses Induced by Strain-Peening Upon Fatigue Strength," by R. L. Mattson and J. G. Roberts. From "Internal Stresses and Fatigue in Metals." Pub. by Elsevier Publishing Co., Amsterdam, Holland.

applied alternating loading, there will be a progressive change in the residual stress as the number of cycles increases.

For alternating stresses near the fatigue limit, the residual stress probably does not change much after the first cycle. There is some evidence in the case of SAE 4340 (Fig. 6) that, after about 10^6 cycles, rather marked changes can occur. The data is so scant in this region that no definite statements can be made concerning this behavior.

With the data and interpretation offered above, we are now ready to proceed to the question, "How should the designer account for changes of residual stresses?"

How Should the Designer Account for Change of Residual Stresses?

This question can be answered only partially at present.

Estimate of Fatigue Limit of Part — Residual stresses should be looked upon as a portion of the stresses which the surface material must resist. These stresses will not change in hard materials with moderately low residual stresses for amplitude of alternating stresses near the fatigue limit. Even in a soft material there will be little change when the member is stressed near the fatigue limit. Hard materials are more sensitive to mean stress than are soft materials. This can be shown by Fig. 19 which is data on SAE 4340 from Morrow and Sinclair.[9] A change of residual stress would occur only in materials which are relatively insensitive to the change.

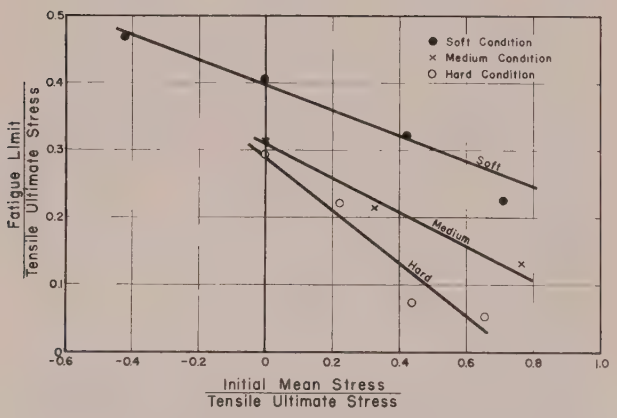


Fig. 19 — Effect of initial mean stress on fatigue limit for SAE 4340 steel

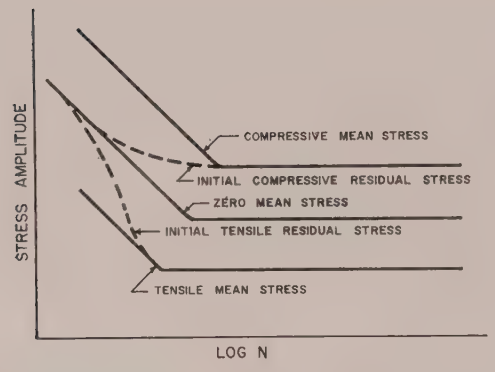


Fig. 20 — Schematic *S-N* curves for mean stress held constant and residual stresses allowed to change

A reasonable estimate can be made of the fatigue limit of a member or part which has moderately low initial residual stresses, in the following manner:

1. Consider the *initial* surface residual stresses to be superimposed on the service stresses.
2. Consider these stresses to remain unchanged during fatigue loading.
3. Treat the problem as any other in which a mean stress is present.

For high initial residual stresses or a load pattern which produces a stress greater than the yield strength of the material at the surface, use Eq. 1 to determine the residual stress present at the end of the first cycle. Consider that the residual stresses do not change after the first cycle, and proceed to estimate the fatigue limit as before. An example of this procedure will be found in a paper by Felgar.[17]

Fatigue Life — For stresses above the fatigue limit there is almost certain to be a change of the residual stresses. With present knowledge, it would be difficult to take this into account. However, one thing can be said: For sufficiently high alternating stresses or for soft metals, the surface material experiences completely reversed stressing after the first few cycles. Initial residual stresses, therefore, have very little effect. An estimate of the life can be made using a completely reversed (zero mean stress) *S-N* curve for the material. Fig. 20 shows the futility of initial residual stresses for large amplitudes in soft materials.

This same figure shows diagramatically how the life would be affected by change in residual stress due to fatigue loading in the region between the fatigue limit and very high stress amplitudes. The fatigue life for this region will be between the life for zero mean stress and the life for a mean stress which is maintained throughout the life of the part and equal to the residual stress. For stresses near the fatigue limit, the fatigue life will be near to the life for the *S-N* curve in which a mean stress was maintained. For very high stresses, the life will be near to the *S-N* curve for completely reversed conditions. The intermediate stress region between these two extremes is still a zone of uncertainty in which the designer must exercise his judgment.

Acknowledgment

This investigation was conducted in the Department of Theoretical and Applied Mechanics as a part of the work of the Engineering Experiment Station, University of Illinois, in cooperation with the Evendale Plant Laboratory of the General Electric Co. Special acknowledgment is due Prof. T. J. Dolan, head of the Department, and Prof. H. T. Corten for their criticism and helpful suggestions. Acknowledgment is due Mrs. N. E. Dahl and R. Swendsen for help in preparation of the manuscript. The assistance and helpful discussion of J. D. Marble and R. D. Halverstadt of the Evendale Plant Laboratory are also gratefully acknowledged.

[17] "Effect of Shot-Peening on Fatigue Strength," by R. P. Felgar. ASME Paper No. 58-SA-46, 1958.

Rating Transmissions from Highway Requirements and Vehicle Specifications

F. C. Meldola

Chrysler Corp

THE RATING of a truck transmission depends on the following variables:

1. Highway requirements.
2. Gross vehicle weight or gross combination weight.
3. Engine torque.
4. Rear axle ratio.
5. Tire size.
6. Number of ratios in transmission.
7. Spacing of ratios.
8. The use of a 2-speed axle or auxiliary transmission.

Since variables 1–5 appear directly in the gradeability formula, and 6–8 are results represented by its usage with the various gear ratios, the gradeability formula can be used as the basic means for rating the transmission:

$$G.A. = \frac{T \times R_t \times R_r \times E \times 1200}{r \times \text{gvw}} - L \qquad (1)$$

where:

$G.A.$ = % gradeability
R_t = Transmission ratio
R_r = Rear axle ratio
E = Drive line efficiency (taken as 0.9)
r = Tire rolling radious, in.
gvw = Gross vehicle or combination weight
L = Rolling resistance in % grade (taken as 1.25)
T = Average applied engine torque, lb-ft Average of maximum net torque and net torque at governed engine speed)

The gradeabilities in the various gear ratios can be calculated for any vehicle specification. By correlating them with a highway requirement probability curve, the per cent of time in each ratio can be obtained.

From a number of test runs over various highways with different tractor-trailer combinations, the time in each transmission ration was clocked and expressed as a percentage of the total trip time. These figures were then matched to the calculated net gradeabilities for the gear ratios in each vehicle. The data, per cent grade versus per cent time in each gear, applied only to the specifications of the test vehicles. As shown in Fig. 1, there is little or no correlation in the plot for the various test vehicles.

By starting with 100% and subtracting successively the per cent time in each gear ratio for each test vehicle, a continuous curve showing per cent gradeability versus per cent time more gradeability required is obtained.[1]

Fig. 2 shows such a probability curve for a normal highway, express turnpike, and rather severe high-

THE GRADEABILITY formula can be used as the basic means for rating a truck transmission. By correlating the gradeabilities in the various gear ratios with a highway requirement probability curve, the per cent of time in each ratio can be obtained. The required hours of gear life for each ratio are then determined, and compared with the available gear life in the ratios.

This procedure gives a detailed analysis of a transmission rating for one vehicle specification at a specified mileage between overhauls. A limitation of the system is that it cannot be applied quickly to various vehicle specifications. The paper outlines the method for constructing a nomogram to overcome this.*

* Paper presented at SAE Summer Meeting, Atlantic City, June 17, 1959.

[1] "Drive Line for High-Speed Truck Engines," by W. P. Mitchell. SAE Transactions, Vol. 62, 1954, pp. 396–408.

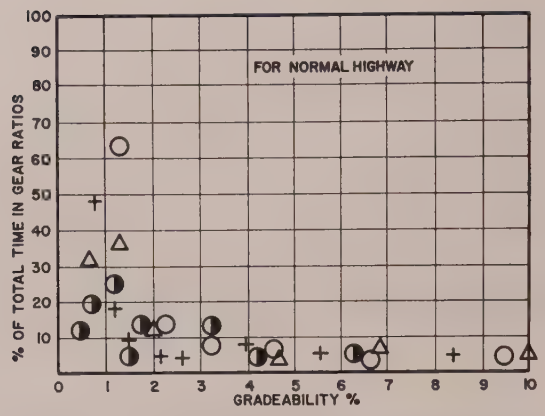


Fig. 1 — Gradeability versus per cent of total time in gear ratios

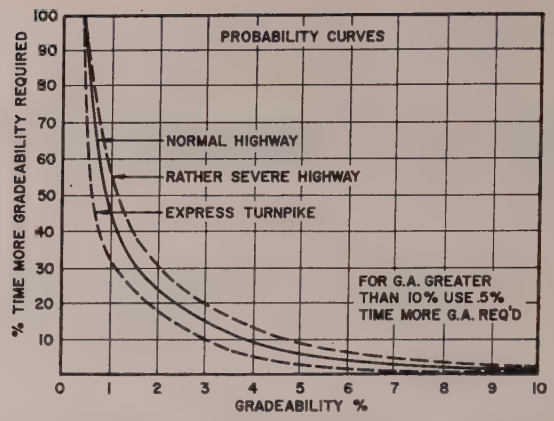


Fig. 2 — Gradeability versus per cent of time more gradeability required

way. This is a universal curve for truck applications. We can change the engine, tire size, transmission, axle ratio, or load, and predict the time in each gear ratio.

Having calculated the gradeabilities in all gear ratios for a specific set of vehicle specifications, we can determine the per cent more gradeability required for each ratio from Fig. 2. Again starting with 100% and subtracting successively, the per cent time more gradeability required for the various gear ratios, we obtain the per cent of total time in each ratio.

The average vehicle speed for each ratio may be obtained from:

$$\text{mph}_a = \frac{\text{rpm} \times r}{168 \times R_r} \times \frac{1}{2}\left(\frac{1}{R_t'} + \frac{1}{R_t''}\right) \qquad (2)$$

The average engine speed in each ratio, therefore, is:

$$\text{rpm}_a = \frac{\text{mph}_a \times 168 \times R_r \times R_t'}{r} \qquad (3)$$

where:

mph_a = Average highway speed
R_t' = Transmission ratio for required average speed
R_t'' = Next numerically larger ratio
rpm = Maximum governed engine speed

Taking an arbitrary figure of 100,000 miles of gear life between transmission overhauls, the total of the time in all ratios may be expressed as:

$$T = 100{,}000/(A \times B,\ 1\text{st} + A' \times B',\ 2\text{nd} + A'' \times B'',\ 3\text{rd} + \ldots \text{etc.}) \qquad (4)$$

where:

T = Total time in ratios, hr
A = Per cent of total time in each ratio
B = Average vehicle speed in each ratio
1st, 2nd, 3rd = Indicators of gear ratios (not used in equation)

The required hours of gear life for each ratio is total time times per cent of total time in each ratio.

This must be compared with the available gear life in the ratios. Two types of gear tooth failure are considered:

1. Failure of the gear teeth by fatigue, usually evidenced by a progressive crack and ultimate breakage at the root of the tooth. The American Gear Manufacturers Association method (Report 101.2) is used. Fig. 3 shows their graph for nonshotpeened, and shotpeened automotive-type gears in terms of cycles to failure versus calculated gear tooth stress. Using the average engine torque and average engine rpm, the available gear life in hours can be obtained for each ratio; also for the main drive and driven countershaft gears.

2. Failure by compressive stress on the gear teeth surface (pitting) is shown by the curve on Fig. 4 in terms of cycles to failure versus calculated compressive stress.[2]

Average applied engine torque is used to determine the calculated compressive stress in the following formula:

$$S_c = 0.59 \cos\psi \sqrt{\frac{6 \times 10^7 T}{f \sin^2 \phi_n}\left(\frac{1}{D_1} + \frac{1}{D_2}\right)} \qquad (5)$$

where:

S_c = Compressive stress, psi
T = Tangential force, lb
f = Width of contact, in.
ϕ_n = Pressure angle, operating normal plane
D_1, D_2 = Respective operating pitch diameter, in.
ψ = Helix angle, at operating diameter

The average engine speed for each ratio is used with the cycles to failure to determine the available hours of gear life.

Fig. 5 is a sample work sheet for a 4-speed transmission. The required hours of gear life per ratio are compared to the available hours for gear failure by both fatigue and compressive stress of the teeth (pitting). The required hours for any one ratio must be equal to the available hours in the corresponding ratio, and all other required hours must be equal to, or less than, the available hours for the corresponding ratios.

When failure by compressive stress is a ruling factor, use the driving gear only for comparison.[2]

If the above conditions are not met by trial and error, select a gvw or gcw which will give the proper balance. This will then be the gvw or gcw rating of the transmission for one vehicle specification.

When a 2-speed axle or auxiliary transmission is used, the low range is treated as another set of transmission ratios as shown in Fig. 5.

While the above procedure gives a detailed analysis of a transmission rating for one vehicle specifica-

[2] "Compressive Failures in Transmission Gearing," by G. E. Huffaker. SAE Transactions, Vol. 68, 1960, pp. 53–59.

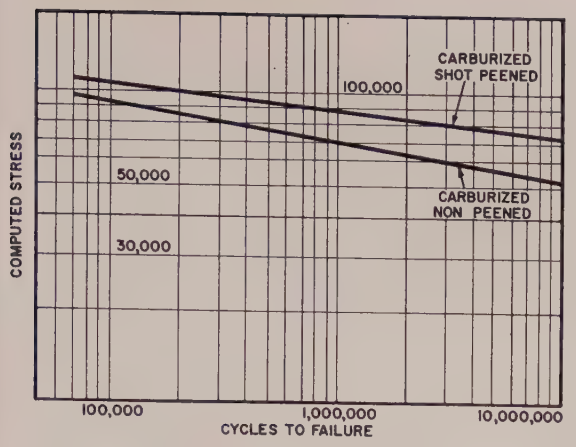


Fig. 3 — Cycles to failure versus computed stress

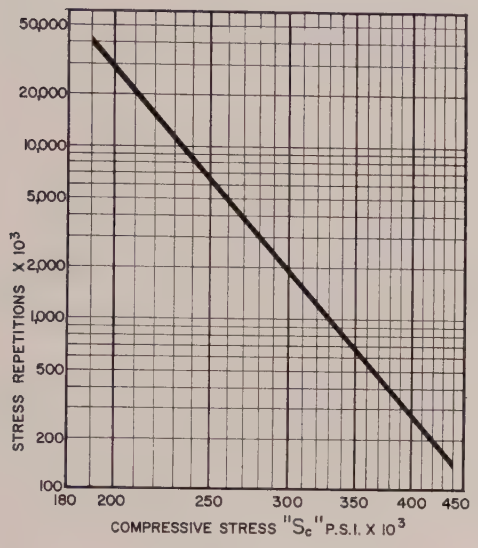


Fig. 4 — Transmission gear life — surface fatigue, stress repetitions versus compressive stress for driving gears

G.V.W. OR G.C.W. RATING — 34,000 LBS.
TIRE SIZE 10-22.5 AXLE RATIO 6.5/9.04

GEAR	NET GRADE ABILITY	% TIME MORE G.A. REQ'D	A % OF TOTAL TIME IN RATIOS	B AVER. M.P.H.	REQ'D HRS. PER RATIO	AVER. TORQUE LB. FT.	AVER. ENG. R.P.M.	⊕ AVAILABLE "C" CALC. GEAR LIFE-HRS. M/S GEAR FAT'G	COMP.	⊕ C/S GEAR FAT'G	COMP.
4H	1.45	32	68	58.5	1350	250	3320				
4L	2.50	18.7	13.3	44.7	263		3550				
4H+4L				→	1613		3485	∞	∞	∞	∞
3H	3.27	12.6	6.1	34.8	120		3330				
3L	5.05	5.5	7.1	25.4	140		3400				
3H+3L				→	260		3365	1330	178	1690	260
2H	7.25	2.1	3.4	18.5	67		3340				
2L	10.55	.5	1.6	13.0	32		3240				
2H+2L				→	99		3290	1210	164	1230	127
1H	16.25	TO START		9.05			3330				
1L	22.95	ON HILLS		6.1			3150	73		70	
1H+1L				→			3240		12		34
MAINDRIVE & C/S DRIVEN GEARS = SUM OF HRS. IN ALL GEAR RATIOS EXCEPT DIRECT					359		3298	∞	372	∞	905
					COMPARE			↑	↑	↑	↑

⊕ USE DRIVE GEARS ONLY FOR COMPARION WITH REQ'D HRS.

Fig. 5 — Work sheet for rating at 100,000 miles of gear life

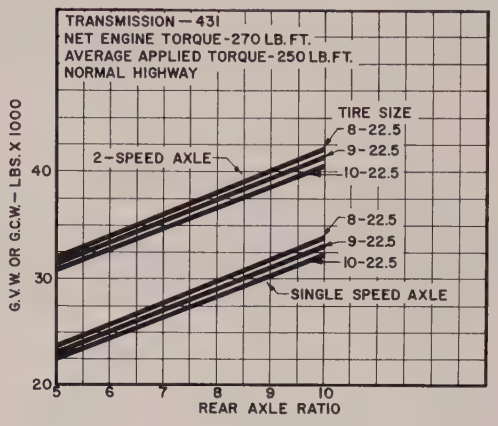


Fig. 6 — Rear axle ratio versus gvw or gcw

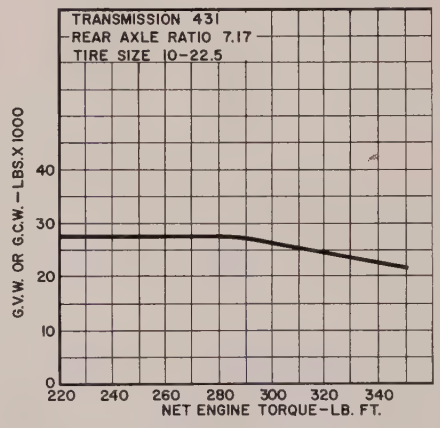


Fig. 7 — Net engine torque versus gvw or gcw

tion at a specified mileage between overhauls, it does not lend itself to fast application of various vehicle specifications. To remedy this condition, a nomogram may be constructed from the results of a few calculations for well-chosen vehicle specifications.

Since the basic probability curves (Fig. 2) are from test results, they do not represent a simple mathematical curve. It is suggested that the nomogram be constructed from calculated curves which box in the practical limits of vehicle specifications. The following curves are required:

1. Fig. 6 shows gvw or gcw versus rear axle ratio. This curve is calculated at one average applied engine torque and one tire size for three different axle ratios. For a 2-speed rear axle or auxiliary transmission, use the high range ratio to obtain the points. Since variation in tire sizes is the same as changing rear axle ratio, plot in lines for other tire sizes by using the ratio of the rolling radii for the basic tire and other tires. For larger tires, the curve will be displaced to the right, and for smaller tires to the left of the basic tire size.

2. Fig. 7 illustrates gvw or gcw versus maximum net engine torque. For one rear axle ratio and one tire size, calculate the gvw or gcw for several average applied engine torques. As the maximum net engine torque is commonly used in engine specifications, plot the gvw or gcw against them since they follow the average applied engine torque closely.

By examining this curve, it will be noted that there is a wide range of engine torques for which there is no noticeable increase or decrease in gvw or gcw of transmission rating. This may be explained by the fact that as the engine size or torque increases, the stress on the gear teeth increases. However, the performance of the vehicle increases. Less time is

spent in the lower gears and more time is allotted to direct drive where the gear life is considered infinite. This balance exists to a point where the increased stress on the gear teeth overcomes the improvement in vehicle performance. There is, then, a rapid degradation of rating in gvw or gcw for increased torque. The transmission becomes less efficient in terms of gvw rating per lb-ft of engine torque. The torque at this turning point may be considered as an indicator of the nominal torque rating of the transmission.

The foregoing discussion has been based on an arbitrary transmission life of 100,000 miles between overhauls. To obtain the rating for any other life period, the curve on Fig. 8 must be constructed.

For any average applied engine torque, tire size, and rear axle ratio, several arbitrary gross weights are selected. The work sheet in Fig. 5 is calculated through items A, B, and C. In the formula for the total time in the ratios, substitute M = miles of life for 100,000 and let H = total time in ratios. Then:

$$H = M/(A \times B, \text{1st} + A' \times B', \text{2nd} + A'' \times B'', \text{3rd} + \ldots \text{and so forth}) \quad (6)$$

where:

H = Calculated gear life divided by per cent of total time in ratio = C/A.

Determine H for the ratios. Use the lowest figure and determine M.

For 2-speed axles or auxilary transmissions, item A is the sum of the per cent of total time in the ratios for high and low range.

The curve as shown need be worked out only once. It is one of a family of curves, all having the same slope. One selected curve is used to establish the variable life line on a nomogram for any transmission.

Using the curves of Figs. 6–8, a nomogram may be constructed as shown in Fig. 9.

For any vehicle specification, for example, maximum net engine torque, tire size, and rear axle ratio, the rating of a particular transmission may be obtained in gvw or gcw for 100,000 miles of gear life, for normal highway operation. By holding the point on the gvw or gcw scale for 100,000 miles of gear life, the gear life for any other gvw or gcw may be obtained as increased or decreased miles of life from the gear life scale.

To construct the nomogram, the line spacing is arbitrary. Spacing in proportion to that shown in Fig. 9 has proved to be very satisfactory.

1. Lay in single speed rear axle scale at any convenient equal spacing.

2. From Fig. 6, use the maximum net torque. This represents the average net torque which was used to calculate the curve. Place a point on the nomogram engine maximum net torque line and label it. From this point, using a straight edge, strike off a point at any convenient position on the reference line and a corresponding point on the tire line. Label the point on the tire line with the basic tire size.

Using the basic tire size on the curve, select a gvw or gcw and its corresponding ratio. Draw a line from the point on the reference line through the axle ratio scale at the proper ratio. Where it intersects, the gvw or gcw rating scale defines a rating. Continue this process until the gvw or gcw rating scale is complete.

For other tire sizes, select a gvw or gcw and its corresponding ratio. Line up these points with a straight edge and strike off a point on the reference line. Line up this point with the original point on the maximum net torque line. Strike off a point on the tire line and label it with the tire size.

3. From Fig. 7 select maximum net engine torque and corresponding gvw or gcw for the axle ratio and tire size. Line up a gvw and the axle ratio. Strike off a point on the reference curve. Line this point up with the tire size and strike off a point on the maximum net torque line. Continue this process until the engine maximum net torque line is completed.

4. From Fig. 8 construct the any other gvw or gcw scale by making divisions equal and in line with those on the gvw or gcw rating scale for 100,000 miles. Invert the scale figures as shown.

Select the gvw or gcw at 100,000 miles of life. Place a straight edge at this point on the gcw rating line for 100,000 miles and the same gcw of the any other gvw or gcw line. Strike off a point on the gear life line and label it 100,000 miles. Retain the same point on the gvw or gcw rating line for 100,000 miles; select other gvw or gcw for the any other gvw or gcw line and mark off the corresponding miles of gear life on the gear life line.

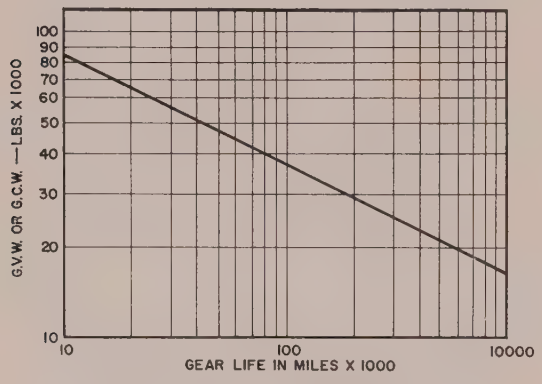


Fig. 8 — Gear life versus gross vehicle or combination weight

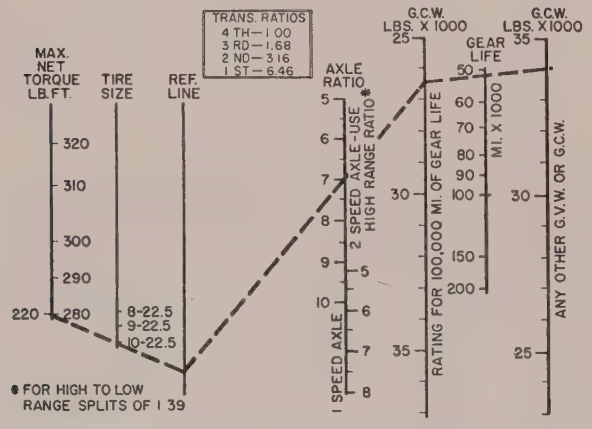


Fig. 9 — Rating nomogram for normal highway operation

Discussion of this paper will be found on pp. 57–59, following the Huffaker paper.

THIS PAPER is a discussion of compressive loading versus stress repetitions resulting in surface failures of transmission gears in laboratory tests. It is also an attempt to utilize data from such tests as a basis for future gear design.

A compressive loading failure is a breakdown of the gear tooth contacting surface by pitting until the surface becomes too rough for proper operation of the gear. This surface breakdown is caused by high compressive loading of the gear tooth surface for many tooth contacts or stress repetition. This surface failure is a surface fatiguing of the metal.

As the fatigue bending life of gears has been lengthened, the incidence of surface fatigue has increased.*

Compressive Failures in Transmission Gearing

G. E. Huffaker
New Process Gear Division, Chrysler Corp.

IN ORDER to design the most economical automotive transmissions, it is important to know how gears fail and what their expected life will be. Failures of hardened steel gears are usually one of three types: (1) scoring, (2) bending fatigue, or (3) compressive loading or surface fatigue.

Scoring, or galling, is a type of failure due to the surface of mating gear teeth welding together in minute local areas. This condition is caused by lubrication breaking down due to high surface pressure, high sliding velocities, and high temperatures.

Bending fatigue failures are caused by the gear tooth being stressed many times. Each time the tooth is stressed, it bends slightly; and after many repetitions of stressing and bending, the tooth breaks off.

A compressive loading failure is a breakdown of the gear tooth contacting surface by pitting, or spalling, until the surface becomes too rough for proper operation of the gear. The surface breakdown is caused by high compressive loading of the gear tooth surface for many tooth contacts or stress repetition. This surface failure is a surface fatiguing of the metal. Fig. 1 is a photograph of a gear which failed in this manner.

Unless gear design is very poorly done, scoring is not a problem in manual-type automotive transmission gearing. Fatigue bending failures and compressive loading failures are items that must be considered carefully.

In recent years, the life of gears in fatigue bending has been increased tremendously, as a result of full fillet root radius gear design, and shotpeening of the fillet root in manufacture. Due to increased fatigue bending life, many gears which previously have failed in fatigue bending are now living longer — but are failing by surface fatigue.

This paper is a discussion of compressive loading versus stress repetitions resulting in surface failures of transmission gears in laboratory tests. It is also an attempt to utilize data from such tests as a basis for future gear design.

Formulas, or equations, for the determination of compressive stresses in gear teeth are usually based on the work done by Heinrich Hertz.[1] Others have

* Paper presented at SAE Summer Meeting, Atlantic City, June 17, 1959.
[1] "On Contact of Solid Elastic Bodies," by Heinrich Hertz. *Journal of Mathematics*, Vol. 92, 1881, pp. 156–171.

elaborated on Hertz' work, and Timoshenko[2] indicates the compressive stress at the center of the contact area between two rolls of different roller diameter but of the same material to be represented by the equation:

$$S_c = 0.59\sqrt{\frac{FE}{L}\left(\frac{1}{d_1}+\frac{1}{d_2}\right)} \qquad (1)$$

S_c = Compressive stress
F = Total load between rollers
E = Modulus or elasticity
L = Length of roller in contact
d_1 and d_2 = Respective diameters of contacting rollers

Fig. 2 illustrates this equation.

The analogy of gear teeth to rolls is represented by Fig. 3. This illustrates how gear teeth may be assumed to have contacting curved surfaces similar to that of rolls. The equation for compressive stress in spur gear teeth may be obtained from the equation for contacting rolls (Eq. 1) and the above analogy to be:

$$S_c = 0.59\sqrt{\frac{TE}{f\cos\theta}\left(\frac{1}{D_1\sin\theta}+\frac{1}{D_2\sin\theta}\right)} \qquad (2)$$

T = Tangential force, lb
f = Face contact width of gear tooth, in.
θ = Operating pressure angle
D_1 = Operating diameter of pinion, in.
D_2 = Operating diameter of gear, in.
E = Modulus of elasticity = 30,000,000

This equation or some similar form has been used by many people to compute compressive stresses in spur gears.

As an involute curve is of constantly changing radius of curvature, the computed stress will be different at each position on the curve. If the terminology of the equation is adhered to, the compressive stress will be determined at the operating pitch diameter of the gear. We believe this to be the proper position for computations as only one tooth will be carrying the load. Test results also indicate most pitting begins near the pitch diameter.

In order to find compressive stresses on helical gear teeth, the radius of curvature of the gear tooth must be determined in a plane normal to the helix. The derivation of a compressive stress equation for helical gears from the spur gear equation is herewith presented. Beginning with Eq. 2 for spur gears, the following items should be added to the terminology listed:

d_a = Equivalent diameter of pinion, in.
d_b = Equivalent diameter of gear, in.
θ_n = Operating pressure angle in normal plane
ψ = Helix angle at operating pitch diameter
f_n = Length of tooth face contact along helix
T_n = Tangential force normal to helix

As the section normal to the helix is an ellipse, we may use an equation of an ellipse and obtain a diameter which would compare to a spur gear of equivalent diameter.

$$d_a = \frac{D_1}{\cos^2\psi} \text{ and } d_b = \frac{D_2}{\cos^2\psi}$$

Then:

$$S_c = 0.59\sqrt{\frac{T_nE}{f_n\cos\theta_n}\left(\frac{1}{d_a\sin\theta_n}+\frac{1}{d_b\sin\theta_n}\right)}$$

$$= 0.59\sqrt{\frac{T_nE}{f_n\cos\theta_n\sin\theta_n}\left(\frac{1}{d_a}+\frac{1}{d_b}\right)}$$

$$= 0.59\sqrt{\frac{2T_nE}{f_n\sin 2\theta_n}\left(\frac{1}{d_a}+\frac{1}{d_b}\right)}$$

$$= 0.59\sqrt{\frac{2T_nE}{f_n\sin 2\theta_n}\left(\frac{\cos^2\psi}{D_1}+\frac{\cos^2\psi}{D_2}\right)}$$

$$= 0.59\sqrt{\frac{2T_nE\cos^2\psi}{f_n\sin 2\theta_n}\left(\frac{1}{D_1}+\frac{1}{D_2}\right)}$$

The normal tangential force $T_n = \frac{T}{\cos\psi}$:

$$S_c = 0.59\sqrt{\frac{2TE\cos^2\psi}{f_n\sin 2\theta_n\cos\psi}\left(\frac{1}{D_1}+\frac{1}{D_2}\right)}$$

$$= 0.59\sqrt{\frac{2TE\cos\psi}{f_n\sin 2\theta_n}\left(\frac{1}{D_1}+\frac{1}{D_2}\right)}$$

Tooth contact length along helix $f_n = \frac{f}{\cos\psi}$:

$$S_c = 0.59\sqrt{\frac{2TE\cos^2\psi}{f\sin 2\theta_n}\left(\frac{1}{D_1}+\frac{1}{D_2}\right)}$$

This equation may be simplified to give the final equation for helical gears.

$$S_c = 0.59\cos\psi\sqrt{\frac{2TE}{f\sin 2\theta_n}\left(\frac{1}{D_1}+\frac{1}{D_2}\right)}$$

Substituting 30,000,000 for the modulus of elasticity E of the material, the equation becomes:

$$S_c = 0.59\cos\psi\sqrt{\frac{6\times10^7\,T}{f\sin 2\theta_n}\left(\frac{1}{D_1}+\frac{1}{D_2}\right)} \qquad (3)$$

As compressive loading failures are surface fatigue failures, such gear failures may be plotted on an S-N type curve. Stress repetition or tooth contacts to failure may be plotted against compressive stresses.

Very little information on compressive loading and resulting gear life is available. The International Nickel Co. has done some work to evaluate gear life. Others have determined safe compressive stresses for specific application. Still others have

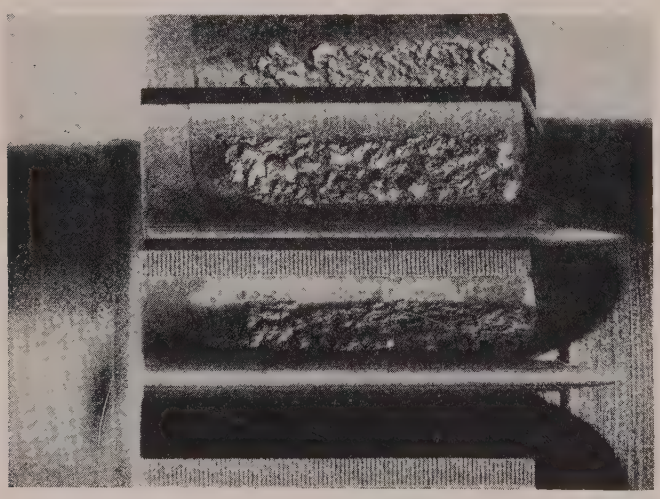


Fig. 1 — Severe pitting

[2] "Strength of Materials," by S. Timoshenko. Part II, pp. 358–359. Pub. by D. Van Nostrand Co., New York.

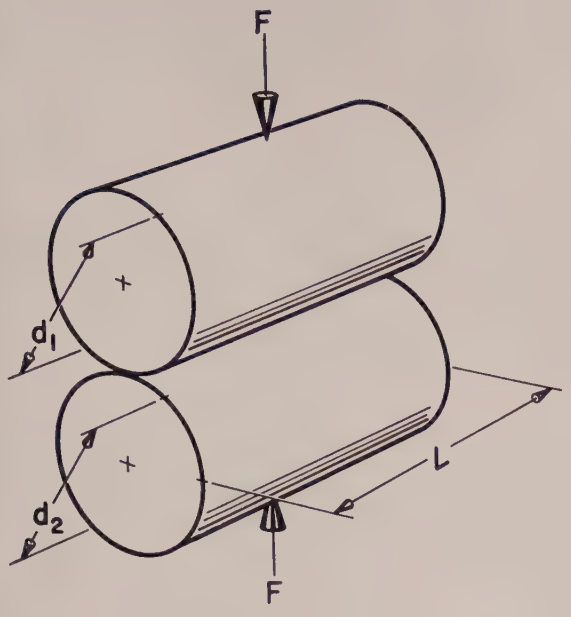


Fig. 2 — Contracting rolls

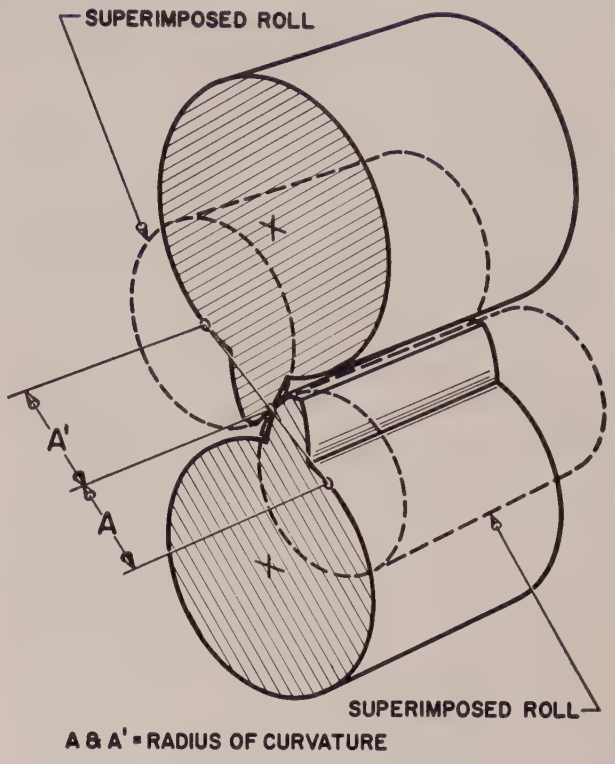


Fig. 3 — Analogy — contacting gear teeth to contacting rolls

suggested compressive stresses based on tests of contacting rolls. In order to have such gear life data available for design use, we decided to attempt to establish a surface fatigue gear life curve including a distribution factor of the type used by the bearing manufacturers.

Bearings fail by surface fatigue similar to compressive loading failure of gears. The bearing industry has been successful in determining usable bearing design data by experimentally running many bearings to failure. Another example of an S-N type curve which was established by experimental testing is described in A.G.M.A. Progress Report.[3] This curve represents fatigue bending and may be used for design work.

Points for our curve were obtained by dynamometer testing of automotive transmissions to gear failure by pitting. A gear was considered as failed when the tooth surface had pitted and spalled to the point the gearset became excessively noisy.

Two dynamometers were used. One was driven by an electric motor and the second by two gasoline engines in tandem. Both setups used electric dynamometers to absorb the output power. Points from either dynamometer drive method corroborate. All dynamometer testing was done at 2200 rpm input to the transmission.

Several models of transmissions were used varying in input torque rating from 200–375 lb-ft. Gears were both spur and helical made of SAE 4023, 4024, 4027, 4427, or 8620 steel, carburized and hardened. The case depth was 0.035–0.045 in. and surface hardness was R_c 58–63. The gears represented production manufacture and were not selected for quality. Involute, lead, and red line charts were made for each gear in addition to the regular production inspection.

For the many points necessary to establish the curve, some tests were made for the primary objective of plotting surface fatigue failures. In addition, all transmission dynamometer testing in our experimental laboratory was observed for surface gear failures. Many of our points came from these routine dynamometer tests.

The gear failure points are plotted on a log-log scale, plotting stress repetition against compressive stress. Compressive stresses were computed using Eq. 2 for spur gears, and Eq. 3 for helical gears. From the hours dynamometer life of the gear and revolutions per minute, the number of stress repetitions per gear tooth were calculated and the resulting point plotted.

Fig. 4 shows the plotted gear failures. Two lines are placed through the points to represent the extremes of gear life. The lower line is placed through the plotted points of gears which failed earliest. The higher line is placed through the points representing gears of maximum life.

For future use of the plotted data in design, it was necessary to establish a working curve, or base line, between the two extremes of gear life. Final working curve Fig. 5 was established by a count of plotted points to give 80% of gear life greater than, and 20% less than the base line. Using the terminology of the bearing industry, this curve would be known as a B-20 gear life expectancy curve.

Equation and Curve

Many designers use speed factors, or dynamic loading factors to modify gear loading formulas. Compressive stress Eqs. 2 and 3 do not have such factors; as a result, the plotted points and final curve (Fig. 5) have no stress load corrections for

[3] "Bending Stress of Spur and Helical Gears." Section I, Progress Report No. 101.02 by Automotive Gearing Committee, American Gear Manufacturers Assoc., October, 1951.

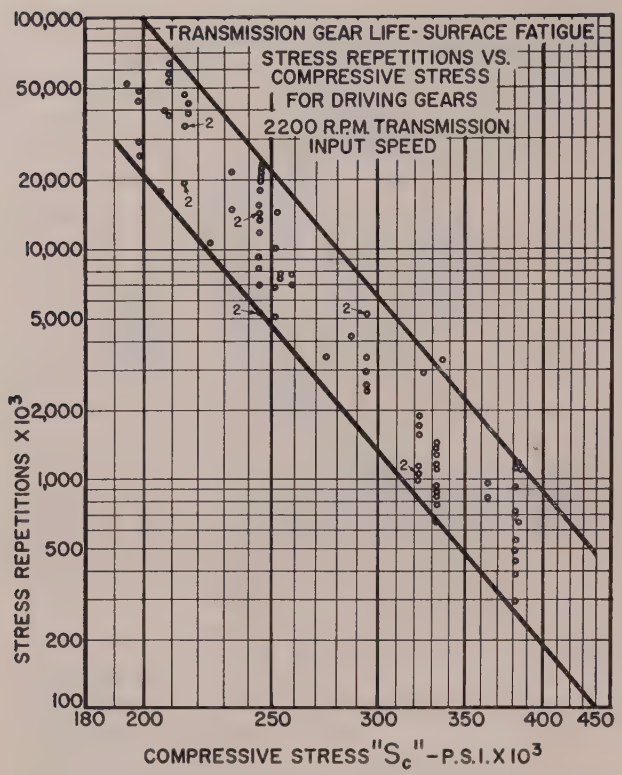


Fig. 4 — Transmission gear life — surface fatigue, stress repetitions versus compressive stress for driving gears (2200 rpm transmission input speed)

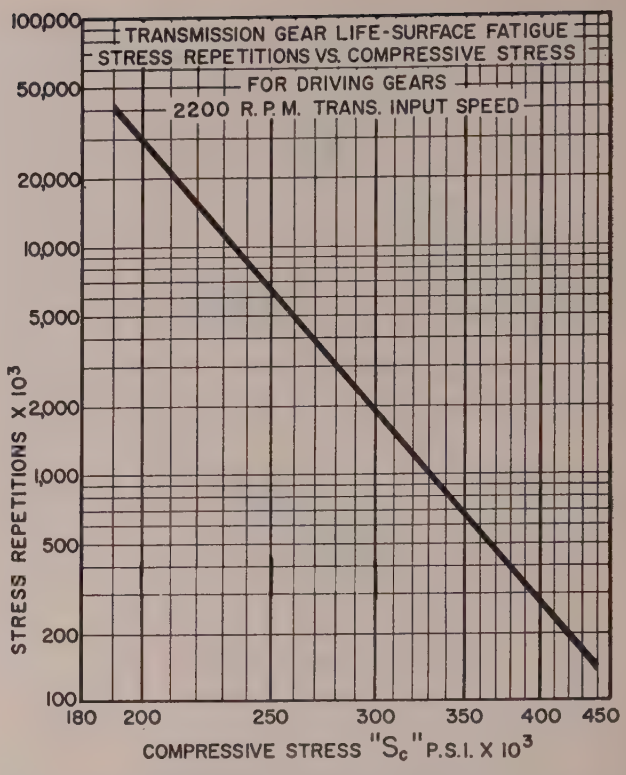


Fig. 5 — Transmission gear life — surface fatigue, stress repetitions versus compressive stress for driving gears (2200 rpm transmission input speed)

the speed of gears. Points, which determined the higher cycle portion of the curve at the least compressive stresses, are transmission input drive gears. Points which determined the lowest cycle gear life and highest compressive loading are first or low speed gearsets. Between these two extremes are second-, third-, and fourth-speed gearsets. With transmission input speeds of 2200 rpm, pitch line velocities vary from 1825 fpm for the input drive gear to 570 fpm for the first-speed gear.

If the curve is used in transmission gear design based on an input shaft speed of 2200 rpm, we believe it has a "built-in" speed factor. Future input drive gears will fall on the same part of the curve as for those tested. Future first-, second-, third-, and fourth-speed gears will fall on a portion of the curve which was obtained by gearsets operating at similar speeds.

If gears that will operate at speeds considerably different than those tested are to be designed, a speed factor should be used. One of our current projects is the determination of such a speed factor by dynamometer testing.

In cases of speed-up ratios where the driving gear member is turning at slower rpm than the output gear, our testing indicated the driving gear failed prior to the driven gear even though the driver had fewer gear tooth stress repetition than the driven. This has been observed by others. D. W. Dudley of General Electric Co. indicates he has performed tests on gear speed-up sets of 4/1 ratio in which the driving gear pitted before the driven pinion.[4] Due to these results, all gear stress repetitions are calculated and plotted for the driving gear member.

In many fatigue curves, where stress cycles are plotted against stress, the curve will bend toward infinite repetitions as the stress is reduced beyond a certain point. We had suspected this might happen near the high cycle end of our curve, as many authorities have indicated pitting will not be encountered at compressive stresses of 200,000 psi and less. Our plotted gear failures did not give an indication of infinite life within the range of compressive stresses tested.

Earle Buckingham has described tests made with steel rolls to define the surface endurance limit of case hardened steel.[5] Tests of up to 400,000,000 stress cycles indicated no definite endurance limit. The load limit followed the same line as for tests of heavier loads and fewer cycles. This would indicate our curve could be extended toward high repetitions and low stress with no bend toward infinite life being encountered.

We would like to explore this high repetitions low stress area, but have not for the following reasons:

1. It is not necessary as most of our design work falls in the area of the defined curve.

2. A great amount of dynamometer time would be necessary to obtain a sufficient number of points as each point in this area of the curve would require many hours of testing. Some of our plotted points represent 500 hr of dynamometer testing.

As loading equations for bearings and gears are both results of Hertz' work on elastic bodies, it is of interest to compare results of testing each. From the literature on bearing life published by the various companies, we find life is inversely proportional

[4] "Practical Gear Design," by D. W. Dudley, pp. 288–289. Pub. by McGraw-Hill Book Co., Inc., New York, 1954.

[5] "Analytical Mechanics of Gears," by Earle Buckingham, pp. 521–522. Pub. by McGraw-Hill Book Co., Inc., New York, 1949.

to the load to a power varying from the third to fourth. Expressed as an equation:

$$\text{Life} = K\frac{1}{(\text{Load})^x} \qquad (4)$$

where:

K = Constant of proportionality
x varies in range from 3–4

Some bearing companies use the third power, others the fourth, and the 3.33 power is often used for roller bearings.

Setting our results in similar form, we find in terms of stress and cycles, the cycles vary inversely as the compressive stress to the 6.75 power, or:

$$\text{Cycles} = K\ \frac{1}{(S_c)^{6.75}}$$

As the compressive stress is proportional to the square root, or one-half power of the load, (Eq. 3) the cycles or life is inversely proportional to the load to the $\frac{6.75}{2}$ or 3.38 power:

$$\text{Cycles} = K\ \frac{1}{(\text{Load})^{3.38}}$$

This 3.38 power is in the range of results found by the various bearing companies by test, and is very close to the 3.33 power which is often used for roller bearings. We believe this tends to prove the analogy of gear teeth to contacting rolls.

Upon examining Eq. 3, it will be noticed the compressive stress may be varied by changing the following gear dimension or items: (1) face width, (2) pressure angle, and (3) helix angle.

In the past, gear designers regulated compressive stress only by changing the face width. Pressure angles were somewhat standardized and helix angles were determined for desired overlap or permissible end thrusts. The value of careful consideration of pressure angles, and helix angles for compact high capacity gearsets, is now known and utilized in gear design.

This paper represents work which is only beginning. Dynamometer testing is now under way to establish speed or dynamic loading factors which will increase the flexibility of the present curve for design use.

Use of Equation and Curve

Assume a truck transmission with a 21-tooth input drive gear mating with a 46-tooth countershaft driven gear. The second speed gearset consists of a 20-tooth countershaft gear and a 37-tooth mainshaft driven gear. The transmission will be operated at 300 lb-ft input torque at 2200-rpm input shaft speed. Find the expected dynamometer life of the second-speed gearset in surface fatigue.

ψ = Helix angle at operating diameter = 33.5 deg
D_1 = Operating diameter of 20-tooth countershaft gear = 3.3333
D_2 = Operating diameter of 37-tooth mainshaft gear = 6.1666
θ_n = Normal operating pressure angle = 16.884 deg
f = Face contact of mating gears = 0.940
T = Tangential force, plane rotation, lb

From Eq. 3:

$$T = \frac{\text{Torque input} \times 12 \times \text{gear ratio driving gearset}}{\text{Radius of countershaft driving gear}}$$

$$= \frac{300 \times 12}{1.6666} \times \frac{46}{21} = 4732 \text{ lb} \qquad (5)$$

$$S_c = 0.59 \times 0.8339 \sqrt{\frac{6 \times 10^7 \times 4732}{0.940 \times 0.5558}\left(\frac{1}{3.3333} + \frac{1}{6.1666}\right)}$$

$$= 246{,}576 \text{ psi}$$

Entering curve Fig. 5 with this compressive stress value, the stress repetition or tooth contacts of the 20-tooth countershaft gear equals 7,300,000.

As each gear tooth is stressed once per revolution, the hours expected life may be obtained as follows:

$$\text{Rpm of countershaft} = 2200 \times \frac{21}{46} = 1004 \text{ rpm or stress repetition per min}$$

$$\text{Life hr} = \frac{\text{Stress repetitions}}{\text{Stress repetitions per min} \times 60}$$

$$= \frac{7{,}300{,}000}{1004 \times 60} = 121 \text{ hr}$$

DISCUSSION OF MELDOLA AND HUFFAKER PAPERS

(Mr. Meldola's paper will be found on pp. 49–52.)

Accurate Probability Curves Needed for Performance Prediction

— Charles M. Perkins
Fuller Manufacturing Co.

MR. HUFFAKER makes use of a method of compressive stress calculation that has been commonly accepted. However, it does not include a speed factor which would normally result in a higher stress figure. Since the final conclusions are derived from actual tests this may not be significant for the gearing that supplied the test results, but a speed factor should probably be included in any formula which attempted to predict the performance of gearing used at speeds other than those that obtained in the tests.

It has not been our experience that the use of helical teeth in place of spur results in appreciably increased life. To determine whether any considerable gain can be expected it would be necessary to run much more extensive tests than the ones that form the basis for this paper. Field experience indicates that straight spur gears would compare very favorably with helical form the point of view of compressive failure, the chief value of the latter being to minimize noise.

Mr. Meldola uses the results of the first to establish a method of rating transmissions on the basis of gear life, taking account of engine torque, speed, axle ratio, gross load, and terrain.

The probability curve (Fig. 2) illustrates the importance of having adequate power-to-weight ratio for top-direct transmissions. The steep slope of the curves at low gradeability shows the marked decrease in the use of gears that will result with a slight increase of horsepower. This is not true in the case of overdrive transmissions, where a slight surplus of power will lead to excessive use of the drive and overdrive gear sets.

The chief difficulty in applying the method here set forth lies in establishing accurate probability curves. The scatter

of the points in Fig. 1 is rather wide and the general pattern does not conform very closely to any of the curves of Fig. 2. A mean curve drawn through the points shown in Fig. 1 would indicate an even steeper slope at low gradeability than is shown in Fig. 2. Assuming that accurate probability curves can be obtained, the method offers a concise means of predicting gear performance, taking account of all the application factors. A similar method applied to the bearings would possibly give a more complete picture of overall transmission performance.

Gears Are Not the Only Troublesome Transmission Part

— W. S. Coleman
White Motor Co.

AS A REPRESENTATIVE of a vehicle manufacturer which purchases its transmissions, I am pleased to comment on these papers which indicate considerable theoretical and experimental study of the design and application of truck transmissions.

Mr. Huffaker's observation that the driving tooth breaks down before the driven tooth, regardless of the comparative gear sizes, leads me to believe that there is a flaw in the theory on which the stress formula is based or on the method in which the formula is applied. The phenomenon appears to have been accepted but not explained. Studying the tooth action, it will be noted that the tooth loading of the driving gear begins near the base of the tooth and progresses toward the tip, while on the driven gear the reverse situation occurs. Possibly, there is someone who can elaborate on the apparent failure of laboratory experience to agree with the theoretical calculation. It would appear that the stress value must be affected by the direction in which the plastic deformation takes place.

Mr. Meldola's method of rating transmissions is based on sound engineering principles, and we at White Motor hope to compare this rating method with our actual experience. It is possible that with similar application of the rating method, other truck manufacturers can accumulate sufficient data to make this a real working tool.

Looking at transmission performance from the operators standpoint, I am sure it becomes evident that there are other phases of the transmission design and production which are, if anything, more troublesome than the gears. Service reports indicate problems such as the following:

1. Slipping out of gear.
2. First speed gear teeth ends chipped off, producing metal debris which gets into the bearings and makes them noisy.
3. Seizure of constant mesh gears on the mainshaft.
4. Tooth failures showing end loadings that may have been used by excessive shaft deflection.
5. Bearing lubrication failures at one end of the transmission and leakage past the oil seals at the other end, due to lubricant distribution within the transmission case.

Extended Probability Curve Would Be More Realistic

— E. E. Eaton
Clark Equipment Co.

FOR MANY YEARS the transmission industry has had the need for some reliable and accepted method of determining transmission life in a given vehicle. The method presented here by Mr. Meldola is a scholarly approach for the solution of this old problem. Transmission life is influenced by a great many variables, and sometimes unknowns. The Meldola paper takes more of these factors into account than any other previously published method and, therefore, has the greatest hope for survival.

Because of the inclusive character of this proposed rating system, the required calculations are comparatively long and tedious. However, if this method of rating is found to be reliable, then the length of calculations required to make the final nomogram for any specific transmission will certainly be warranted.

Upon trying this method during the past year by calculating many known transmission experiences in highway trucks, it appears that probability curve (Fig. 2) should be extended and used beyond the 10% gradeability value. If the instructions are followed to use 0.5% more gradeability required for all conditions over the 10% gradeability, the resulting required hours for certain gear ratios may not be compatible with known experience.

To illustrate, referring again to Fig. 2, the probability curve for normal highway is extended beyond the 10% point. Then, by using readings from this curve for conditions of greater than 10% grade, calculations are made covering a specific transmission installation for the required hours per ratio and set down as suggested in Fig. 5 of the paper. It is noted that with the extended curve method, the required hours for 100,000-mile life in our specific example is 59 for second gear and 8 for first gear, while the Meldola proposal gives a corresponding figure of 57 and 0, respectively. These latter figures would indicate that the first gear is not used; therefore, a 4-speed transmission rather than a 5-speed unit should be satisfactory. Yet, the truck industry says that five speeds are required. Thus, it appears that figures obtained from the extended curve are more realistic.

At the outset of this discussion, we indicated that transmission life is influenced by many variables, all of which makes an accurate mathematical analysis difficult. Some of these variables appear as assumptions in the formulas of the paper, such as: average torque, operating schedule, average vehicle speed, drive line efficiency, and the rest. Other factors not mentioned are also involved, such as lubrication, vibrations, product quality — not the least of these is the influence of the man operating the gearshift lever.

Transmission Test: Vehicle-Test Course Compatibility

— J. H. Letsinger
International Harvester Co.

MR. MELDOLA'S paper is an excellent application of basic principles and calculation procedures and has prompted us to take a further look at our testing procedures and providing ground results.

By observing a test vehicle on our 7¼-mile Phoenix Proving Ground test course and recording total time in gear (sec) and total engine revolutions in gear, we can plot the vehicle's compatibility with our test course as a transmission test. Calculations needed are per cent time more gradeability required, average rpm in gear, net torque, and per cent gradeability.

Fig. A shows the calculated performance of an R-225 on the 6% grade portion of our test course superimposed on the probability curves used in Mr. Meldola's paper. Results to date bear out the compatibility shown.

Fig. B covers an ACF-195D on the 9% portion of our test course. The transmission in this vehicle is a 5-speed with auxiliary-type box and has five consecutive speeds in high range and five consecutive speeds in low range. You will note that the top five speeds by their relation to the curves indicate light-to-average usage while the three low speeds shown indicate severe-to-very-severe service. The auxiliary mainshaft bearing failed at 50,000 miles of service.

Fig. C covers an A-170 on the 6% portion of our test course. We have shown this to point out the lack of compatibility. We did not need the figure to know this was a poor transmission test.

In all the figures, we have been unable to show the lowest gear run in as it plots as zero with regard to per cent time more gradeability required and gives no measure of severity of service.

We believe this application of Mr. Meldola's approach to transmission design will enable us to quickly evaluate our future tests and allow adjustment of one or more of the test variables to secure the type test we wish to run.

Mr. Huffaker's Closure To Discussion

AS MR. COLEMAN has questioned the reason for the driving gear of a gearset failing before the driven gear, I would like to offer the following explanation. I cite as a reference, D. W. Dudley.[a]

The tooth loading, as Mr. Coleman indicated, begins near the base of the tooth and progresses to the tip for the driving gear; the opposite being true for a driven gear. During the progression of a tooth contact, both driving and driven gear teeth are going from sliding at first contact to rolling at the pitch line to sliding in the opposite direction as the teeth leave contact.

The sliding action for the driving gear tooth is away from the pitch line toward the root of the gear during the first part of contact and away from the pitch line toward the tip of the gear during the last part of the contact. For the driven gear the reverse is true and the sliding is from the tip toward the pitch line and from the root toward the pitch line.

When a high compressive load is applied to the tooth, small minute surface cracks develop. If no other force was present the cracks would develop normal to the contacting surfaces. As the metal is pushed or pulled by the sliding action, the direction of the progression of the cracks is influenced.

For the driving gear the cracks develop from the surface with a directional component toward the pitch line. For the driven gear the cracks tend to develop in a direction away from the pitch line. Thus, the cracks on the driving gear develop toward intersection, and on the driven gear, they develop away from intersection. (See Fig. D.) When the cracks have developed to intersection, the surface concerned will spall and flake away.

We believe the above explanation indicates it is not contrary to theory to use laboratory results of **driving** gear life and calculated compressive stresses to establish gear life performance curves for driving gears.

ORAL DISCUSSION

— Reported by R. E. Fletcher
Dana Corp.

C. Kope, Ford Motor Co.: How do you account for the difference in torque between maximum net and governed net?

Mr. Meldola: The curves are based on an average applied net torque.

R. Kaufman, Chevrolet Motor Division, General Motors Corp.: Does crown shaving have an effect on the compressive stress curve?

Mr. Huffaker: The compresive stress curve is based on gears having the larger gear in the set crown shaved. Extreme crown shaving would increase the compressive stress.

H. W. Christenson, Allison Division, General Motors Corp.: Transmission life is determined by the minimum life in the critical gear in the transmission. How can you make this system meet various customer requirements?

Mr. Meldola: With this system, the transmission can be rated for any type of operation. It would be the responsibility of the selling organization to control the operation and recognize unusual conditions.

M. Schall, Dana Corp.: Has any allowance been made for using transmission behind torque converters, or behind diesel engines with their characteristic vibration?

Mr. Meldola: The tests reported in these papers were not run behind diesel engines or converters, so no answer is available.

G. R. Beardsley, Ford Motor Co.: Were any variables added to compensate for engines being 4, 6, or 8 cylinders?

Mr. Meldola: Little or no difference was noted in constructing the nomogram for either 6- or 8-cyl engines of equal horsepower output.

[a] "Practical Gear Design," by D. W. Dudley, pp. 288–289. Pub. by McGraw-Hill Book Co., Inc., New York, 1954.

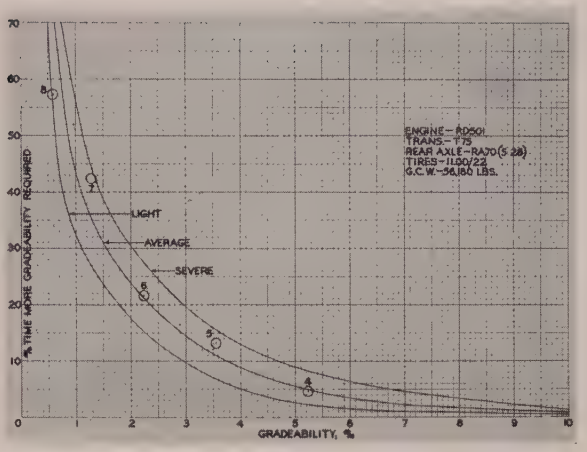


Fig. A — R225 on 6% test course

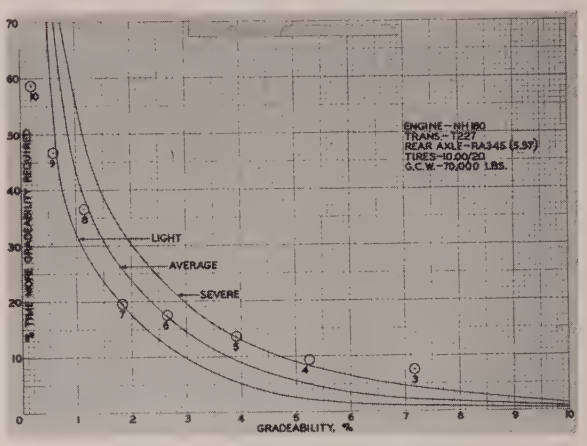


Fig. B — ACF195D on 9% test course

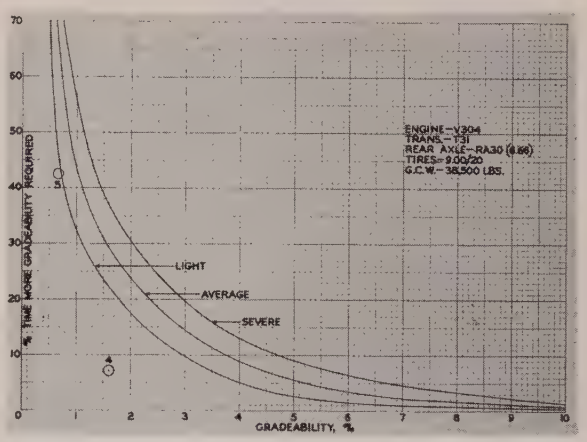


Fig. C — A170 on 6% test course

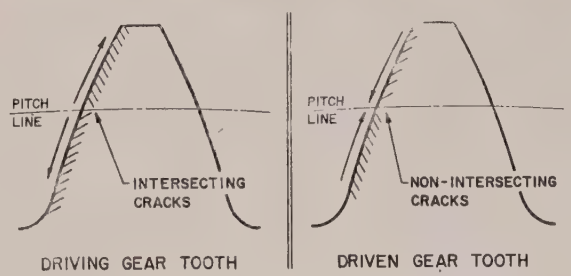


Fig. D — Development of minute surface cracks (direction of sliding indicated by arrows)

a study of Self-Contained for Turbojet and

Henry R. Schmider and John H. Ferguson, Jr.
Bendix Aviation Corp.

THE TURBOJET and turboprop engines of today require much more energy for starting than did the reciprocating engines with which we are all familiar. As modern jet engines are increased in size and complexity, this difference in energy requirement becomes even more evident.

As a comparison, the reciprocating engine normally starts when rotated only a few revolutions at relatively low speed and requires only small amounts of energy. Starting of turbojet engines, on the other hand, requires acceleration of the engine rotating parts to speeds of about 30–35% of normal rated speed for satisfactory starting, while turboprop engines often require acceleration to speeds in excess of 60% of rated speed. Acceleration to speeds lower than these values generally results in "hot" starts which can cause serious engine damage, or at least reduction in engine operating life.

Because of the nature of the jet engine, wherein a compressor pumps air into the turbine section to provide combustion air, the load on the starter is actually increasing during engine acceleration prior to engine light-off. The reverse is true in the case of the reciprocating engine, where overcoming the inertia of the sliding parts is performed in the first moments of rotation, with resultant decrease in the torque requirement from that point until the start is accomplished.

It is not uncommon for the kinetic energy of the rotating parts of a large turbojet engine of the present day to be in the order of 1,000,000 ft-lb at the completion of the starting cycle. This is the equivalent of a 4000-lb automobile traveling at a rate of 87 mph. While not quite all of this energy is imparted by the starting equipment, as the engine provides some assistance after it reaches self-sustaining speed, the bulk of the energy is provided by the starting equipment. Conversely, the kinetic energy in the rotating parts of the reciprocating engine is very low.

Thus, it is apparent that substantial power is necessary to start the modern jet engine.

While the early and relatively small jet engines could be started with electric starting equipment similar to, but generally much larger than that used for reciprocating engines, new methods for starting jet aircraft engines were needed.

Turbine-driven starting equipment appeared to provide the most advantageous means of developing the torques necessary to crank the larger series of jet engines through the starting cycle. The first practical system developed consisted of two major pieces of equipment: the gas generator for providing the energy source and the starter drive unit, consisting of the turbine, reduction gearing, clutch, and engine engaging mechanism. In these early starting systems, the starter drive unit was mounted in the airframe, while the gas generator was a piece of mobile ground equipment connected temporarily to the starter drive unit during the engine starting cycle and disconnected before taxiing.

The push button start remembered from reciprocating engine days was sorely missed and the need for starting systems completely contained within the airframe was apparent. Such assemblies could be made to produce considerably higher starting torques, starting the engine faster, while eliminating the need for ground equipment connection and disconnection. Actually, the aircraft could be made

Starting Systems for Turboprop Engines

airborne at least 1½ min sooner than the ground-supported systems, possibly the difference between a successful or futile intercept mission.

Also, self-contained starting systems would allow operation from bases not equipped with the ground carts necessary for a ground-supported system. The harassment of unavoidable landings at bases not equipped with ground handling equipment was certainly a major factor leading toward the search for suitable self-contained starting equipment.

The basis for such equipment, the starter drive unit, had at least been established in concept. The problems that remained centered mainly around the development of suitable gas generators to provide the motive power for driving the turbine. These gas generators necessarily had to be lightweight, durable, and reliable.

This paper is concerned only with the equipment which is used in one way or another with self-contained starting systems. Since in some cases there is an overlapping of equipment necessary for this with equipment not exclusively applicable to self-contained systems, some discussion has been included where deemed advisable to complete this presentation.

History

The desire for self-contained starting systems originated with the military personnel. It was they who had the earliest requirements for such equipment because of the need for fast operational capability and widely scattered service usage. If suitable systems could be developed, unlimited potential usage could be foreseen.

Aircraft accessory manufacturers began work leading toward development of self-contained starting systems employing almost every conceivable method of gas generation. Foremost among these methods were:

1. Combustion of a fuel-air mixture
2. Burning of a solid monopropellant
3. Decomposition of a liquid monopropellant

Steady progress was made in this country on fuel-air combustion and solid monopropellant starters. Generally, the widest acceptance for liquid mono-

SUBSTANTIAL POWER is necessary to start the modern jet engine. Thus, starting equipment has become a major concern of air transport operators. This paper discusses the equipment used with self-contained starting systems.

The authors discuss and evaluate a variety of self-contained sytsems: combustor, fuel-air combustion, cartridge, liquid propellant, hydraulic supported by auxiliary power units, and electric supported by APU. Possible future systems are: self-breathing systems, oxygen combustors, and liquid-oxygen-water-fuel combustors.

It is emphasized that the choice of a starting system for a particular aircraft will depend on aircraft characteristics and the aircraft's intended use.*

* Paper presented at SAE National Aeronautic Meeting, New York, March 31, 1959.

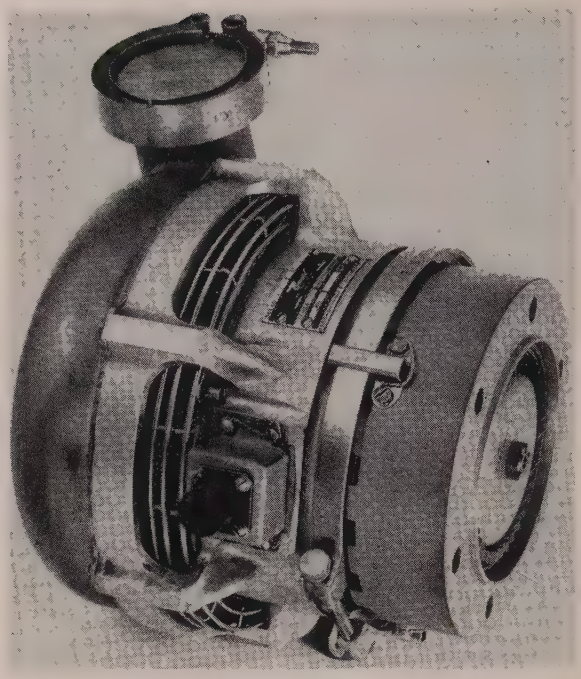


Fig. 1 — Typical air turbine starter

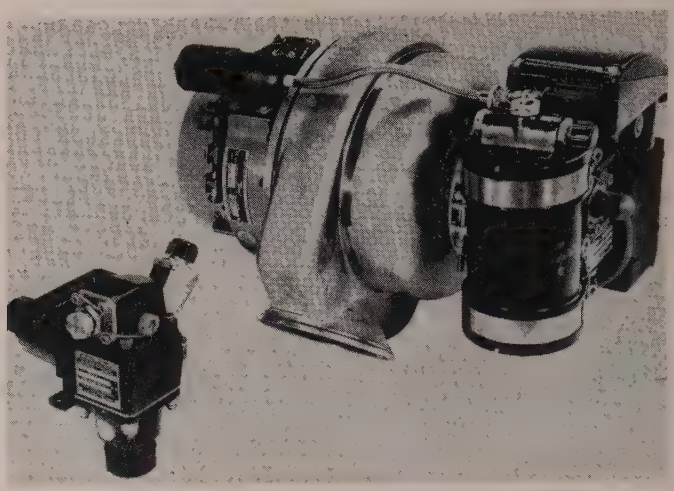


Fig. 2 — Fuel-air combustion jet engine starter

propellant starters was attained in Europe.

Successful completion of fuel-air combustion starters and solid monopropellant starters led to their widespread usage on the J-65 engines for the F84F and the B-57, respectively. Concurrently, fuel-air combustion starters were used for the J-73 engines on the F86H.

Development of solid monopropellant cartridges did not keep pace with the concepts necessary to provide reliability and durability. The first cartridges were smoky, erosive, and produced corrosive products of combustion. In addition, the logistics problems arising from the character of the propellants limited acceptance of the starters requiring them.

For these reasons, the fuel-air combustion starter was more widely accepted for future versions of military aircraft, being specified for the F-102, F-106, F-101, and KC-135, to name a few. All of these aircraft are presently equipped with fuel-air combustion starter systems.

Developments toward improvement of solid propellant cartridges have been reasonably successful, in that the characteristics of the propellants have been improved, and the cost of cartridges reduced somewhat. This has led to recent renewed interest in cartridge starting. This type of self-contained device is presently being considered for several advanced air vehicles and presently is specified for the F-105.

Because of these two suitable means of providing the aircraft in military usage with the advantages of self-contained starting systems, relatively little work has been done on liquid monopropellant systems in the United States. However, systems operating with the energy produced by the decomposition of hydrogen peroxide, propyl nitrate, and hydrazene have been explored and at least partially developed in this country or Europe.

In general, the commercial airlines, looking for a relatively economical and reliable system which could be made self-contained or ground-supported as individually desired, also specified variations of fuel-air combustion starters for their fleets. These systems generally produce gas at energy levels substantially below the military versions of either fuel-air combustion or solid propellant starters and hence require larger mass flow rates of gas for somewhat longer periods of time. Hence, when contained within the air vehicle they are somewhat heavier and occupy more space.

Starter Drive Units for Self-Contained Systems

At this point, a brief description of the starting equipment which is available for use in self-contained systems is in order.

Direct Impingement — This system can provide the simplest form of self-contained starting system. Here, gas produced by some form of generator (products of fuel-air combustion, solid propellant, liquid monopropellant, steam, and the like) is impinged directly on the turbine of the jet engine through separate starting nozzles in a manner which will cause acceleration of the rotating parts through the necessary speed range. Because the gases cannot, in all cases, be impinged against the turbine in the most advantageous direction without the nozzle hardware interfering with the main airflow through the engine, systems of this type are often necessarily quite inefficient. They often require 2–4 times the amount of gas flow which would be necessary with geared turbine starter units especially designed for the purpose.

For this reason, the direct impingement method is presently impractical for large engines. Also, the objections of most engine manufacturers to passage of possibly incompatible gases of questionable origin through the turbine sections of their engines have minimized acceptance of these systems.

Pneumatic Starter — The pneumatic starter generally consists of an air inlet manifold, turbine, reduction gearing, an overrunning clutch or disk clutch, and engine decoupling mechanism. A typical unit is shown in Fig. 1. This unit is mounted either on the engine rotor, engine gearbox, or in

some instances attached to the engine gearing by means of a readily removable probe device during the starting cycle only.

Air, generated by ground cart apparatus, airborne gas turbine compressors, or in the case of multiengine installations, by the compressor stages of an engine previously started by another type of starting device is impinged against the starter turbine through suitable ducting. The temperature of these gases seldom exceeds 500 F at a pressure generally below 50 psia.

There are many operational aircraft using pneumatic starting equipment. Two installation concepts are generally employed on single engine aircraft, either a ground-supported engine-mounted starter, or a probe-type starter also supported from ground sources. In the case of multiengine aircraft, the ducting is often installed so that the engine-mounted starters may be operated from ground sources, or from air bled from the compressors of running engines.

The energy level of air supplied to pneumatic starters can be considered only moderately high. Therefore, the starting cycles are generally longer than the self-contained units which will be discussed.

Because the airborne weight and size of the gas turbine compressors needed to supply air for pneumatic starters is generally prohibitive, pneumatic starting as such is not considered to be within the scope of this presentation. However, these starters can be considered as secondary equipment for multiengine aircraft using a self-contained starter for one or more engines, with compressor cross-bleed air for the remainder. These units have a definite place in systems where airborne weight is an important factor, as compared with rapid simultaneous starting.

Pneumatic Starter with Combustor Support — With a change in the support concept for pneumatic starters, a completely self-contained starting system can result. A starter unit, basically a pneumatic starter, can be supplied with air from a stored source, which for space considerations, is generally pressurized to 3000 psi. The stored air is regulated to a lower pressure, generally about 50 psia, with suitable valving and passed through a combustor. A supply of jet engine fuel is also admitted to the combustor and the mixture of fuel and air is burned to produce a moderately high energy gas in the order of 700–1000 F which is ducted to the starter drive unit.

The amount of air necessary to produce a suitable start may be considerably reduced by the addition of water or some other suitable coolant which serves the purpose of increasing the mass flow of gas through the turbine. Since the torque characteristics of turbines are mainly dependent on mass flow, the addition of coolant results in a substantial saving of air, with a resultant saving of the weight of the air plus the effect of reduced air storage tank size and weight.

Installations using combustor-supported pneumatic starting equipment supported by either ground or airborne carried air have been widely accepted for usage by the airlines to support their jet fleets. The systems used vary considerably in detail from one fleet to the next, with economic considerations for a particular type of operation generally being the governing factors influencing equipment selection.

In order to make aircraft employing the combustor-supported pneumatic starter adaptable to completely self-contained service, an air compressor is added which refills the air storage tank during flight, so that the system is ready for restarting when required. The weight of this type of starting system can generally be lightened by the use of high-pressure fiberglass reinforced plastic storage tanks using the fiberglass winding to provide strength. These tanks generally weigh about 2 lb for each pound of air contained at 3000 psig.

Fuel-Air Combustion Starter — The most widely used self-contained starting system to date is the fuel-air combustion type. This equipment is capable of delivering very high torques for engine acceleration, minimizing the time required for starting and also almost completely eliminating the hot start problems which occur often on hot days when lesser flows of air are supplied to the engine turbine by the compressor.

Here, the gearing and turbine section of the starter drive unit are similar to the equipment for a pneumatic starter in principle except for increased ruggedness and material changes necessitated by the high temperatures of the generated gases. A typical fuel-air combustion starter is shown in Fig. 2.

Gas for acceleration of the starter turbine is produced by combustion of a mixture of aircraft jet engine fuel and relatively high-pressure air (about 200 psig) at mixture ratios controlled by design characteristics from stoichiometric to about 35/1. The resultant gas produced impinges against the starter turbine wheel at a velocity of 4000–5000 fps. This gas velocity provides a very high turbine runaway speed and, consequently, produces a much flatter starter output torque curve than is provided by pneumatic starters which have lower gas velocities, hence low runaway speeds, and steeper torque versus speed curves. This characteristic of the high-energy series of starters provides more usable output torque to accelerate the engine during the starting cycle, without producing higher impact torque which might damage the engine.

In these starters, fuel for the combustion process is taken from the main engine fuel tanks, so that no special fuels are necessary on board the aircraft for the starting system.

In completely self-contained systems, the air necessary for combustion is compressed by a small airborne air compressor usually taking a very small amount (0.15–0.45 lb per min) of bleed air from the engine compressor. This air is stored in high-pressure tanks at 3000 psi, generally in fiberglass wound containers.

In many cases, this air compressor is already on board the aircraft for other purposes. In this case, it is usually desirable to increase the capacity of the air storage system. However, this can only be determined by full investigation of the system.

In cases where complete self-sufficiency is not necessary, the air storage tanks may be filled from ground sources. This, of course, results in lower airborne weight for installations not having compressed air for other purposes. Many operational aircraft have this type of system on board. The advantages of fast engine starting are retained while eliminating the need for ground servicing at one or two stopover points. Of course, the airborne-sup-

ported system, which eliminates all ground support with small additional weight penalty presents the only high-energy completely self-contained system.

Solid Propellant Starter — This again is a high-energy engine starter similar to the fuel-air combustion starter in general design characteristics from the turbine wheel to the engine. However, the gas generation and control system is of completely different concept. The output torque is of generally similar characteristic, in that it produces a relatively flat output torque at varying speed throughout the normal operating range. Fig. 3 shows a photograph of a solid propellant starter for engines of the J-57, J-75 power class.

In the case of the solid propellant starter, the gases for impingement against the starter turbine are produced from the burning of an ammonium nitrate cartridge, which is burned in a chamber which can either be mounted on the starter or located remotely with ducting to the starter unit. The charge is ignited electrically from the cockpit.

The charge burns at a constant rate which is dependent on the temperature of the charge at the time of ignition. The burning rate is such that the entire charge burns at substantially the same rate, as the temperature of the charge does not vary during the cycle even though the gases produced are in the vicinity of 2000 F.

A charge at a low ambient temperature on ignition burns at a slower rate than a high-temperature charge, and hence the starter produces a lower torque for accelerating the engine. A charge at – 65 F produces a starter torque about half that of a charge at + 160 F. This factor introduces a problem in matching the starter to the engine, in that the output torque of the starter at low temperature must be capable of overcoming the engine drag, whereas the output torque at high temperature cannot exceed the engine pad torque limitations. This is an important consideration in setting up requirements for cartridge starters.

Cartridges produced in the early 1950's generated smoky gases of highly erosive and corrosive nature which precluded the production of starters having the necessary degree of reliability and durability.

Recent developments by the cartridge manufacturers have at least partially solved these problems. Further developments by the cartridge manufacturers are necessary, however, to reduce cost, provide cartridges with burning rates substantially unaffected by temperature and clean enough to allow better control of the gases after ignition.

The solid propellant starter systems are as self-contained as the number of cartridges carried on board the aircraft or at least available to the crew. In degree of self-containment, the system is comparable to fuel-air combustion systems without airborne compressors, but having a fixed volume of air stored on board.

Cartridge starters have the advantage of being lighter than a fuel-air combustion starter system having the same capability in number of starts, but have the disadvantages of requiring convenient access to the cartridge chamber for loading and high cost per start.

Solid propellant starters have been made operationally acceptable, and are presently being considered for several new high-performance aircraft.

Liquid Monopropellant Starters — Here a charge of liquid monopropellant is decomposed to produce the high-energy gas needed for turbine impingement. Monopropellants which can be considered suitable include highly concentrated hydrogen peroxide, isopropyl nitrate, and hydrazene. All are difficult materials to handle, and principally because of this there has been no operational installation of such equipment in this country, nor does there appear to be any serious consideration being given to such installations.

They may, however, be considered for installations where the monopropellant fuel is already on board the air vehicle for some other purpose, such as fuel for an auxiliary power unit.

Little difficulty should be experienced in developing such units should they become required, as much experience in the decomposition of the materials has been gained for auxiliary power unit work using all of these fuels.

Hydraulic Starters — Hydraulic equipment can be employed for starting of self-contained aircraft under certain specific conditions.

In this system, a hydraulic pump motor unit is directly connected to the engine gearing. During the starting cycle, the unit acts as hydraulic motor being supplied with a flow of hydraulic fluid. After engine starting, the unit can be made to operate as a pump to provide hydraulic fluid for necessary aircraft functions.

In order to provide self-contained starting, this equipment is generally considered feasible only if a power unit must be operated either as an alert pod or an auxiliary power unit for a prescribed period prior to take-off of the aircraft. This power unit can then be made to supply high-pressure hydraulic fluid for driving the starting motors.

Since some new high-performance aircraft do require such warmup periods for certain equipment, hydraulic starting is being considered for normal starting. Generally, some additional self-contained starting method is being considered for abnormal starting where alert pods are not available.

If an auxiliary power unit is on board, but not used normally for supply of hydraulic power, electrically driven hydraulic pumps can be used to supply the starting motors. This normally results in increased

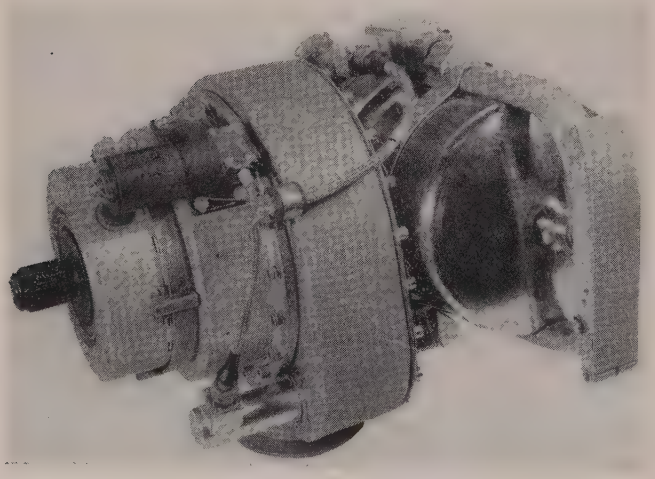


Fig. 3 — Solid propellant jet engine starter

weight chargeable to the system.

Hydraulic starting equipment provides some advantages in presenting no exhaust problems and a relatively high efficiency. On the other hand, however, unless large amounts of hydraulic fluid are needed for aircraft services, multiengine installations will be quite heavy compared to other self-contained installations.

Electric Starters — Electric starting equipment can hardly be considered at this time as adaptable to self-contained starting of large jet engines. This is especially true if the air vehicle is not equipped with some form of auxiliary power unit which supplies electrical power which can be used for starting, because the weight of batteries is prohibitive for this purpose. Starter generator units, however, may be attractive for some installations where part of the unit weight can be charged to power generation after serving as a starter motor. However, the power being provided by the auxiliary power unit added to the additional power being provided from the starter generator units, provides more d-c power than any known installation requires. A-c starting units are quite impractical at the present time.

In the case of the small series of engines which are presently being developed, the electric starting picture is somewhat different. Here, consideration can be given to starting one or more engines of a multiengine installation with electric starter generators either from batteries or auxiliary power units, while others may be started with air bled from compressor stages of running engines. As many starter generators can be used as needed to supply the electrical requirements of the aircraft in flight. If batteries are used, however, the system may be heavier than other self-contained systems, but the high degree of reliability achieved in electric starting may in some cases prove advantageous.

While electric starters may be operated from either battery or controlled current power supplies, the latter is preferable. A given starter unit operating from a controlled current power supply provides greater output power than when operated from batteries. A 1000-amp controlled-current starter will provide a peak output of about 25 hp.

Adaptability

In contrast with the period of exclusive use of the reciprocating engine, there are today many types of suitable engine starting equipment available to those responsible for selection. With so many types of equipment available for installation, selection is very much a matter of customer evaluation. Often, the choice of proper equipment for a particular application is a controversial issue between members of the same organization because equipment providing advantages from one standpoint may be less acceptable from another. Briefly, there is no system available which is best for all installations.

Because of this, no attempt is made here to define the best starting equipment for a jet aircraft of a particular type. An attempt is made here, however, to point out the considerations which must be evaluated by those responsible for the selection. The following items are of major importance, though not listed in order of relative importance for this can only be done when all facts relating to a particular installation are available.

The first step in selection of proper aircraft starting equipment is the evaluation of self-contained versus ground-supported systems. Analysis of this requires the following considerations.

1. Is it important to get the aircraft into the air very quickly?

If it is, one of the high-energy, self-contained starting systems will probably save, on the average, 2–3 min over ground-supported systems when applied to multiengine aircraft, and 1–1½ min on single-engine aircraft. Probably about half of this saving can be realized using a high-energy self-contained system when compared with a moderate or low-energy self-contained system.

If there is no need to become airborne rapidly, a self-contained system is not necessary on the basis of this consideration.

2. Will the aircraft be operated periodically from remote and widely scattered bases?

If so, self-contained starting is almost mandatory. While many bases today are equipped with ground carts, they may not be suitable for a particular installation because of the multitude of different possibilities for ground-supported systems.

Aircraft which may be used for advanced military buildups, where the air vehicles may become "sitting ducks" until ground-support equipment becomes available, generally require self-contained starting systems.

3. What will be the ratio of aircraft to operating bases?

If the number of aircraft is high compared with the number of bases of operation, self-contained starting may not be as economical as ground-supported starting. However, if the ratio is low, self-contained starting systems are often more economical to maintain. If economy is important, as in airline service, this factor must be considered thoroughly. This will be discussed somewhat more thoroughly later.

Even though self-contained starting systems may not be necessary based on any of the above considerations, they may be desired for other reasons which are brought out by investigation of the above. If they are desirable, and are to be specified, investigation based on the following considerations must be undertaken to determine the best system.

Mission or Flight Schedule — Here again, if it is important to save the time between starting engines with high-energy self-contained systems and the lower energy systems, the choice can be narrowed down to fuel-air combustion, solid monopropellant, or liquid monopropellant for all engines of the air vehicle. Using one of these systems on each engine makes it possible to start all engines in 15 sec, and have them operating above idle speed in 20–25 sec.

The length of the mission or flight program is also important in determining the type of system. The lower energy systems, especially the fuel-air combustor generally require more air per start and, consequently, longer storage tank refilling times or necessitate larger air compressors. If the mission is to be short, this may again limit consideration to only the high-energy systems.

Reliability and Durability — Certainly the reliability and durability of the equipment must be a major consideration. More service experience has been gained on fuel-air combustion equipment than any other type of self-sustaining starting equip-

ment. While some difficulties were experienced during development and early service usage, these have been corrected, and today, fuel-air combustion starters probably give a higher proved reliability than any other self-contained system. Considerable work in proving the durability and reliability of cartridge starters and combustor starters is underway at this time, and full evaluation of the status of these programs must be coordinated with the suppliers of starting equipment.

The reliability and durability of equipment is closely related with maintainability of a fleet of aircraft. Starting equipment is often difficult to install and remove. Use of equipment which proves unreliable has a "snowballing" effect in cost of aircraft ground time, overhaul of equipment, and necessary stockpiling of components and units.

Weight of System — The aircraft designer is invariably fighting a weight problem. It is difficult to state which system is lightest until all facts are considered. In any case, the degree of self-containment must be the same in considering the weight. It is unrealistic to compare a self-contained system with airborne recharging capability with a system carrying fuels for a limited number of starts.

Cartridge starters, while generally providing the lightest weight, may not be advantageous where a limited number of starts is undesirable. A cartridge for a modern jet engine when suitably packaged weighs 12–13 lb. The number of cartridges carried plus starter weight will determine the total system weight.

Cost of Operation — While extremely difficult to evaluate, this is of the utmost importance. There can be no doubt that cartridge starting equipment is by far the most expensive to operate (the charges themselves cost about $50).

Logically, cartridge-type starters cost at least 10 times as much to operate as any other type in use.

Direct operating costs of other types of starting equipment vary considerably, but are, in the main, approximately equal and very low. Major costs at the moment stem from the relatively short overhaul life of the equipment and trouble shooting required. However, fuel-air combustion starters have demonstrated 1000-cycle capability between overhaul periods.

Installation Considerations — The physical location of the starting equipment in the airplane is important to certain types of self-contained equipment. For example, the breech of a cartridge starter must be located in a readily accessible location for reloading. The breech can, of course, be located remotely from the starter drive unit with suitable ducting for the gases to pass the starter drive unit.

In the case of other types of turbine-driven starters, the duct connections must also be in an accessible location for removal of the equipment.

Equipment requiring periodic lubrication or checking of lubricant must, if selected, be accessible for this purpose.

Cost of Installation — Initial cost is perhaps a relatively minor consideration. The cost of system installation is relatively unimportant when all factors including cost of operation, cost of overhauls, replacement, and the like, are fully analyzed. Generally, cartridge starters afford the lowest initial cost, but when all factors are considered result in the most expensive installation available today.

Table 1 is intended to give an idea of the relative characteristics of self-contained starting systems.

Combination Starting Systems

In multiengine installations, the following system concepts are quite practical for self-contained starting systems depending on the operational concepts. The following list is based on a four-engine aircraft; however, the list is equally applicable to other multiengine installations in principle:

1. Four fuel-air combustion starters.
2. Four cartridge starters.
3. Four fuel-air combustor starters.
4. One of any above with three air turbine starters.
5. One of above with three hydraulic starters.
6. One of above with three electric starters (some very special installations).
7. Two of above with two air turbine starters.
8. Two of above with two hydraulic starters.
9. Two of above with two electric starters (in some very special installations).

While these seem to be the most logical contenders for consideration, the number of possible combinations is really quite endless.

Combinations of some of the heavier high-energy self-contained units with lighter not normally self-contained devices have the effect of making multiengine installations lighter while retaining more of the advantages of complete self-containment. The combination of equipment is being considered and used on increasing numbers of aircraft.

The major disadvantages result from having different kinds of starting equipment in the supply line, and the loss of economies resulting from smaller purchases of different equipment. However, in the case of systems with relatively high operational costs, overall economies may easily result.

Combination Starter

During recent months there has been an increasing demand for combination starter units, wherein a self-contained starter can also be operated from ground supplied energy sources. In this equipment, the gearing and turbine are generally common for both starting modes, but the gases from two different sources can be made to impinge against the turbine wheel. Generally, only slight increase in weight is necessary to accomplish a dual mode of operation.

Advantages inherent in dual purpose starters allow use of lower cost operational starts from ground sources for normal flight plans, with the less normal start achieved from higher cost starts under the self-contained mode. Major cost advantages exist for cartridge-air turbine combination starters.

In some cases, combination starters have been considered desirable as a back-up operational device in case the engine cannot be started by the first attempted mode. The fallacy of this arrangement stems from the probability that if a deficiency in the starter drive unit exists, the turbine or gear section will be at fault, and since this hardware is common to both modes, the starter will be inoperative under a second mode.

However, for those who desire this dual operational mode, the following combinations are practi-

Table 1 — Self-Contained Starting Systems

Type of System	Combustor	Fuel-Air Combustion	Cartridge	Liquid Propellant	Hydraulic Supported by APU	Electric Supported by APU
Degree of Self-Containment	Can be completely self-contained	Can be completely self-contained	Limited by number of cartridges	Limited by supply of propellant	Can be completely dependent on supply for APU	Can be completely dependent on APU
Weight	Very heavy	Heavy	Medium	Medium	Very high if chargeable to starting system	Very high if chargeable to starting system
Logistics Problem	None	None	Supply of cartridges	Special fuels	None dependent on APU fuel	None dependent on APU fuel
Cost of Installation	High	High	Low	Low to medium	Depends on amount chargeable to starting	Depends on amount chargeable to starting
Cost of Operation	Low	Low	Very high	Can be very high	Low	Low
Reliability	Should ultimately be very good	Proved by many varied installations	Good	Not proved	Should be good, but little service experience as a starter	Excellent
Durability	Should ultimately be very good	Good, with very good chance for immediate improvement	Questionable but new cartridges should increase this	Not proved	Should be good, but little service experience as a starter	Excellent
Remarks	Weight can be reduced by removal of air-borne recharging or by water augmentation for combustion	Weight can be decreased further by removal of recharging system	Possible installation disadvantages because of need for accessibility of breech	May be more desirable if special fuel is available on air vehicle anyway	Probably impractical except in cases where unique features of air vehicle exist	System is probably impractical except for unique installations

cal, and can be produced by the starter manufacturers:

1. Combination cartridge-air turbine.
2. Combination fuel air-combustion-air turbine
3. Combination fuel air-combustor-air turbine

The merits of specifying this equipment can only be appraised by the user.

Dual-Purpose Starters

The lack of geared pads on the engine for mounting equipment, such as generators, pumps, constant speed drives, and the like, has recently become an important problem. For this reason, the elimination of a usable engine pad taken up by a starter is sometimes undesirable.

The starter industry has met this challenge with units capable of dual functions. Here, an auxiliary drive pad is provided on the starter on which a hydraulic pump, generator, and the like, can be mounted. It is logical to mount the accessory on an auxiliary drive pad provided by the starter, rather than the starter on a modified accessory redesigned to provide a drive pad. Because the starter is the higher torque unit, shafting, gearing, and so forth in the accessory would necessarily be strengthened if the starting torques were transmitted through these units.

Units can and have been designed which provide dual purpose within a single package. Potential combinations which may result in weight savings for specific applications are: starter constant speed drive, starter air-conditioning unit, starter generators, starter motor-hydraulic pump, and so forth.

Economics of Operation

Economics of operation of any starter system is important, especially in the competitive position of airlines operation. For this reason, considerable effort was expended during the preparation of this paper to secure comparative data on costs for the various systems which might aid in optimum selection. Because of the very incomplete nature of these studies, mostly due to the small amount of experience gained by the air fleet operators to date, such comparison is impractical at this time. It should, however, provide an interesting study for the future.

In order to ascertain the true cost of starting using any system, all items bearing a relation to the program must be considered. These naturally include, but are not limited to:

1. Initial cost of equipment to operate a given size fleet.
2. Cost of materials used in operation.
3. Overhaul costs.
4. Operator costs.
5. Necessary spare parts and units.
6. Rapidity with which starting can be accomplished which results in: (a) Fuel saving prior to take off, and (b) saving engine and airframe time.
7. Weight which is carried aloft and which is carried as dead weight.
8. Amortization of capital expenditures.
9. Depletion of invested capital.
10. Many more items peculiar to particular operations.

An attempt is made in Fig. 4 to show the type of curve which is desirable in comparing the economics of ground-supported systems and aircraft contained starting systems. Plotted here, without actual values, are the ratio of aircraft to operating bases versus the total cost per aircraft per year for starting systems for both ground-supported and self-contained systems. Generally, the cost figure for self-contained systems will be substantially fixed, regardless of the ratio of aircraft to bases. The

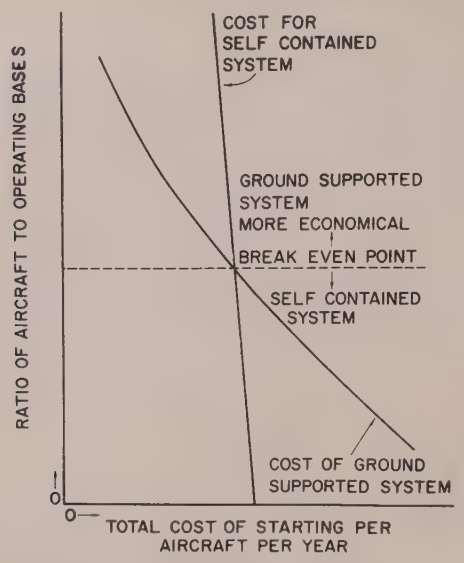


Fig. 4 — Economics of self-contained starter system compared with ground supported systems

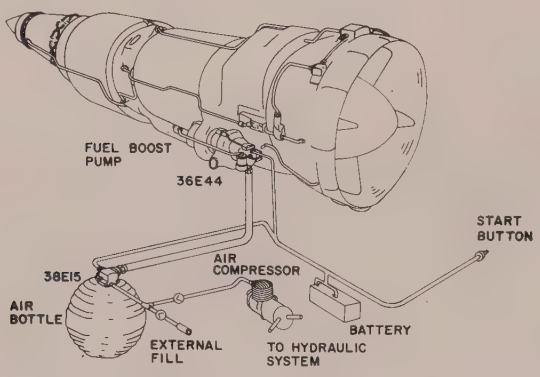


Fig. 5 — Single engine self-contained starting system schematic

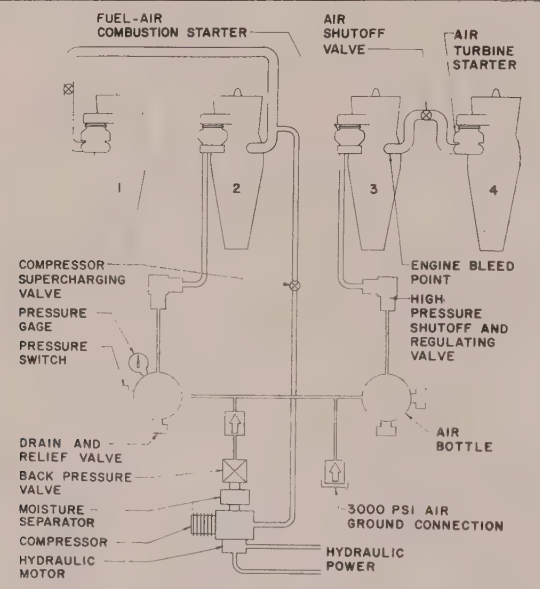


Fig. 6 — Self-contained multi jet engine starter air system schematic

major variations result in slight economies resulting from lowered spares support, reduced training costs for personnel responsible for replacement, and the like, as the ratio increases. The curve for ground-supported systems, on the other hand, will show a drastic change as the ratio changes. Obviously, this is so because the purchase and maintenance of several ground-supported starting systems for very few aircraft becomes prohibitive. As the ratio of aircraft to operating bases increases, the cost of equipment chargeable to each aircraft drops below the relatively fixed cost of self-contained systems. The point at which this cross-over occurs is the break-even point. With ratios higher than this, ground-supported systems result in more economic operation. With ratios lower than this, self-contained systems are more economical.

This plot is especially desirable for fleets which are expected to change in mode of operation over a period of time, for it will show the character of the economics for later additions to the fleet or to the system route plan. In this way, a more intelligent selection can be made in determining the best long-range starting system.

Future Planning

A logical question centers around the future planning of the starter industry for both military and civilian applications. Will a single self-contained system be developed in the foreseeable future which will be universally acceptable for all applications? In short, the answer is probably not. While all starter manufacturers are working on new ideas and concepts in starting which are likely to gain wide acceptance, no single known concept is likely to provide features sufficiently unique to preclude the use of all others.

While each starter manufacturer, because of the competitive nature of the field, guards his ideas and developments closely, some possible future systems can be mentioned briefly:

Self-Breathing Starting Systems — Here, a starter which would in fact be a miniature jet engine, with a compressor to pump air for its own combustion, would be used to accelerate the aircraft engine. This type of equipment has been used by the Russians for some time, but has not found much favor in this country to date. Studies performed to date have shown no particular advantages for this equipment over systems developed in this country.

Starters Employing Oxygen to Support Combustion — In a fuel-air combustion starter, the air used for combustion contains 80% inert nitrogen. Considerably lighter storage systems can be produced if this material can be eliminated by the use of oxygen alone to support combustion. The gas thus produced would have a higher energy than those presently used. The main problem associated with production of such starters stems from the high gas temperatures which result from the combustion process. As new high temperature materials are developed, however, this may become more feasible.

Starters Employing Liquid Oxygen, Water, Fuel Mixtures to Support Combustion — These materials are already found on many aircraft of the present day for other purposes. Little or no increase in the present system capacity would be necessary to provide self-contained starting. Here, a mixture of

liquid oxygen and fuel would be burned to produce a high-energy gas, with water being added in the combustion chamber to cool the mixture and increase the mass flow of gas impinging against the starter turbine.

The major deterrent to development of this equipment is the rather high rate at which liquid oxygen would have to be converted to gas which would necessitate much larger gas converters than are presently necessary for the normal usage rate of the oxygen.

This system may become more practical with any widespread usage of reigniting engines after flameout by filling the combustion chambers with oxygen. This would require piping the liquid oxygen to the vicinity of the engine, making it available for starting, and minimizing the resultant weight penalties.

Conclusions

There are several important conclusions which may be drawn from a study of self-contained engine starting equipment:

1. The importance of self-contained starting systems in starting both military and civilian aircraft has been established.

2. The aircraft accessory industry has met the challenge, and produced a wide variety of components which can be integrated into the required starting systems.

3. No one system can be considered optimum for all aircraft. Considerable study of all aircraft accessory features must be undertaken before the best system for a given installation can be selected. Only by weighing the relative advantages and disadvantages of each with an importance factor assigned to each can the proper selection be made.

4. There are available self-contained starting systems which are light in weight, versatile in characteristics, and which have exhibited a high degree of reliability and durability. Because of the large number of systems in use, both reliability and durability are improving constantly.

5. Starter system analysis is a specialized field when applied to jet engines. The starter must be accurately matched to the engine. Too much torque often results in increased starting time because of engine control considerations. Too little torque may result in increased starting time, or complete failure to start. Because of these tailored requirements, it is invariably desirable to secure advice on engine starting early in the design stages of the aircraft. The starting equipment manufacturers, with years of experience behind them, stand ready to render this service.

Bibliography

Under the Cooperative Engineering Program of the Society of Automotive Engineers, a committee of specialists representing the starter industry and the engine manufacturers has been formed. It is the function of this committee to work closely with representatives of the Armed Services toward the dissemination of information of a relatively noncompetitive nature, as well as toward establishing a degree of unformity in the equipment specifications which establish common denominators for discussion with the users. In this way, standards for certain calculations may be established which eliminate confusion on the part of the users of this equipment.

The committee has published several publications available through SAE headquarters. A list is provided here for information purposes.

AIR #37 Pneumatic Starter Composite Outline Drawings.

AIR #61 Hot Gas Turbine Starter Composite Outline Drawings.

AIR #62 Guide for Determining Engine Starter Pad Torque Requirements.

APPENDIX I

General Starting System Design Parameters

This section covers the general design parameters for starters and starting systems employing turbine-driven starters. Turbine-driven starters are used in the majority of the self-contained systems of aircraft equipped with modern large gas turbine engines. Special aircraft with primary requirements for large amounts of hydraulic or electrical power may be self-contained with the starters using this energy. These are, in the main, considered special cases and will not be covered in this section.

Two very general starting system schematics are presented in Figs. 5 and 6. These schematics serve to point out the major components of a fuel-air combustion system.

The Engine — An aircraft gas turbine engine consists of a compressor, combustion chambers, and a turbine. The turbine serves to drive the compressor and is connected directly to it by means of a shaft. There are many different versions of the basic engine such as the twin spool, turboprop, turboshaft, turbofan, and the like; however, the general characteristics of these engines are the same. Since the general engine characteristics are the same, the starter requirements are also similar. The only differences result from specific design and performance characteristics which determine the magnitude of the torques and speeds during the start cycle.

In starting an aircraft gas turbine engine the starter must provide sufficient power to the engine to overcome three loads. These are:

1. Overcome the compressor drag below engine self-sustaining speed.

2. Accelerate the inertia of the rotating engine mass.

3. Provide power to accelerate the accessories connected to the engine.

The engine compressor must be accelerated to a speed high enough to provide combustion pressure in the engine combustion chambers. After the engine light-off point is reached, the engine turbine begins to supply power to the compressor, and as the speed of the engine increases, it supplies ever increasing amounts of power. When the engine turbine output equals the compressor requirement, a self-sustaining point is reached. Increased engine speed results in excess power output of the engine turbine for further acceleration until the fuel flow is throttled back. A typical steady state characteristic drag-assist curve of a jet engine is shown on Fig. 7.

A starter is required to accelerate the engine up to a speed where it is capable of operating in a satisfactory manner. This speed must be greater than minimum self-sustaining speed and less than the

idle speed of the engine. The starter torque and the engine torque are both shown as dashed lines on Fig. 8.

The sum of the torque resulting in the system for engine acceleration is shown as the net torque on Fig. 8. As long as this net torque is positive, it is available for accelerating the rotating mass of the engine.

The time required to accelerate the engine can be calculated by evaluating the equation:

$$t = \int_0^N \frac{dN}{T_{net}} \qquad (1)$$

where:

T_{net} = Equation for net torque line in terms of N

Since this calculation is time consuming, a very close approximation is often used to calculate this acceleration period. This is done by establishing the net torque available for acceleration for small increments of speed change.

The time to accelerate the engine across each increment is then calculated from the equation:

$$t = \frac{I_e \, \Delta N}{307.1 \, T_{net}} \qquad (2)$$

where the output of the starter is used, or:

$$t = \frac{(I_e + I_s) \, \Delta N}{307.1 \, T_{net}} \qquad (3)$$

where steady state starter torque is used.

The difference is the torque necessary to accelerate the starter itself. It is usual practice to present steady state torque curves for air turbine starters, and output torque for combustion and solid propellant starters.

The sum of these small increments of time quite accurately presents the time to any given engine speed. A typical time versus speed curve is shown as Fig. 9.

Caution must be exercised in selection of the speed increment, as this method assumes that torque is constant across such increment. Improper selection can present an inaccuracy especially at the point where the net torque approaches zero.

In order to permit a comparison of various engines and the relative starter requirements of these engines, a calculation of the total kinetic energy in the engine rotating parts at starter cutoff is useful. Fig. 10 will aid in an approximation of this calculation and will yield numbers with enough accuracy to be useful. Many of the so called "small" engines are not so small when it comes to starting requirements. The inertia may be small but the starter cutoff speed is usually high and the energy varies as the square of the speed.

To determine the energy in a given engine at starter cutoff (Fig. 10), find the point where the engine inertia and starter cutoff speed intersect. The constant energy lines for several levels are drawn to aid in the estimation.

The time required for the engine to accelerate from starter cutoff speed to any speed can be calculated in the same manner except that T_{net} becomes the engine torque only. It is standard practice to present time *through* idle. In this calculation the engine torque curve is extrapolated as if the fuel control did not cause its return to zero torque at the idle speed. If this were not done, the time increments would become very large as the engine torque curve approached zero at idle speed, resulting in a very long and unrealistic time.

Starter Calculations —

1. Torque

The following basic equations can be used to predict the output torque of axial flow impulse-type turbine-driven starters. This type of wheel is presented due to its wide acceptance in starting equipment. There are some starters made with radial flow wheels. A radial wheel results in a lower torque in the low-speed and high-speed range of starter turbine operation. Although use of a radial flow wheel may result in a higher peak efficiency, the overall operation generally results in equal or slightly poorer performance.

The output torque of an impulse turbine can be predicted from the expression:

$$T_T = \frac{\eta R w}{g} \left(C_1 \operatorname{Cos} \alpha + \psi \, \omega_1 \operatorname{Cos} \beta_2 - u \right) \qquad (4)$$

The output torque of a starter can then be de-

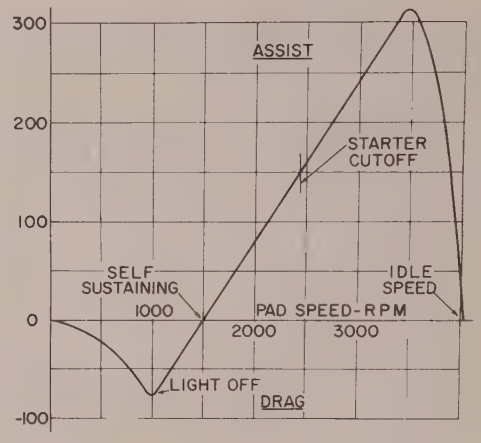


Fig. 7 — Typical turbojet engine steady-state torque versus speed characteristic during start cycle

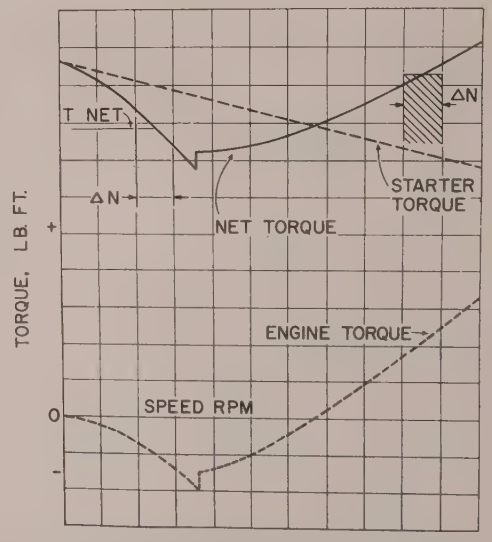


Fig. 8 — Torque versus speed curve for increment method of calculating time

termined from:

$$T_o = \eta_g R_g \times \frac{I_e}{I_s + I_e}(T_T - T_w) \qquad (5)$$

The following special notes are in order:

C_1 — The nozzle velocity can be calculated from standard thermodynamic expressions. In the case of a solid propellant starter it is necessary to correct this velocity for off design conditions due to the under or over expansion losses. Since the nozzles can be designed for only one condition and there is a large change in the operating pressure over the cartridge operating temperature range, this loss can amount to a sizable loss.

I — The engine and starter inertia figures must be taken at the same location. The effective inertia changes inversely as the square of a gear ratio.

T_w — This turbine windage loss can be obtained from empirical data. In the combustion starter or solid propellant starter, it is low to negligible compared to the other figures. In the case of an air turbine starter, it must be determined.

A curve showing the typical output torque for a fuel-air combustion starter and an air turbine starter is shown on Fig. 11. The solid propellant curve at a given temperature would have the same characteristic as the fuel-air combustion starter; however, at different propellant temperatures, the curve would be considerably different in magnitude. Typical horsepower curves are shown on Fig. 12.

2. Speed

The peak speed of a starter turbine should be as fast as possible while meeting the necessary durability requirements. The speed to which a turbine can run is dependent on the material and the operating temperature. The speed of an air turbine starter turbine wheel is limited by gas velocity, and these types of starters have wheels built to operate at their no load speed without destruction. Air turbine starter turbine wheel runaway conditions will be about 1700 fps at the pitch line, with normal cutoff at about 1500 fps.

For solid propellant, combustion and most mono-

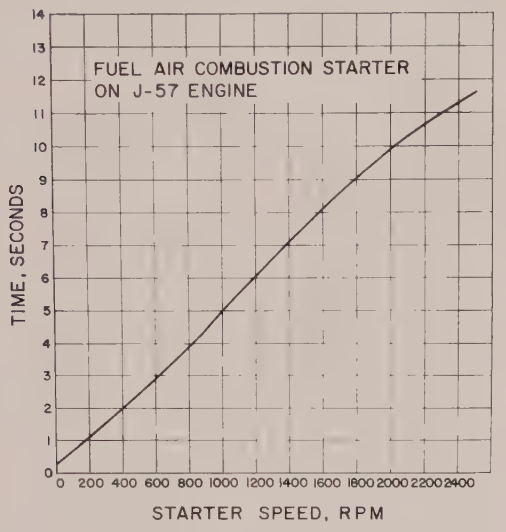


Fig. 9 — Calculated cranking time for typical turbojet engine

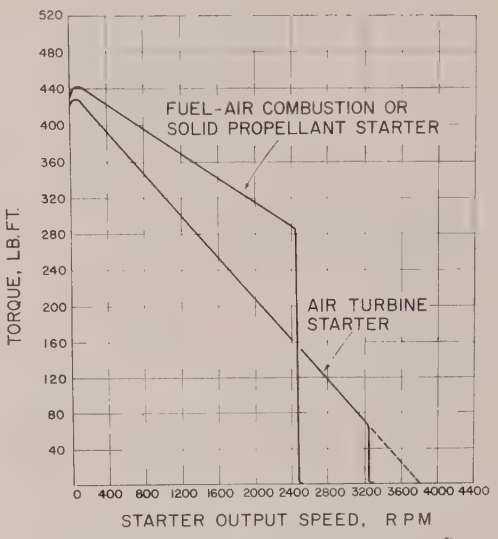


Fig. 11 — Typical starter characteristic curves for turbojet engine starters

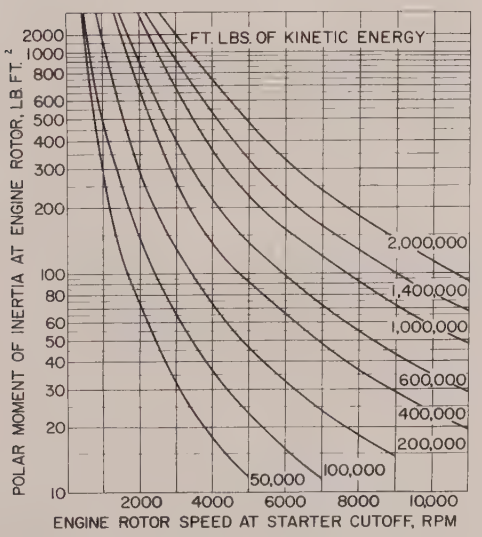


Fig. 10 — Kinetic energy in rotating mass

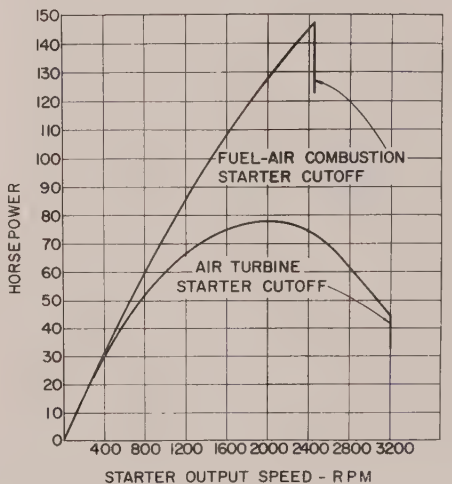


Fig. 12 — Typical horsepower characteristic of starters suitable for 600,000 ft-lb kinetic energy output

propellant starters, the gas velocity exceeds the safe wheel operating speeds. Normal cutoff for these starters is about 1100–1200 fps. For a 6-in. pitch diameter turbine, this amounts to 42,000–46,000 rpm. The starter gear ratio is obtained by dividing the allowable turbine speed by the required engine pad speed at cutoff.

3. Air

The air supply for the air turbine starter is determined from the ground cart or APU output or the engine bleed conditions. A regulating valve is usually used if this air supply pressure is to exceed 50 psig, since turbines are pressure sensitive devices. The stall torque of a given starter is not a function of air temperature, but is very sensitive to pressure changes. The air temperature will affect the output as the speed increases, with higher temperatures providing more output.

The fuel-air combustion starter usually operates from air stored at 3000 psi. The optimum energy output will occur at combustion-chamber pressures of 200–300 psi. Higher pressures will decrease the yield from the air bottle without increasing performance correspondingly and lower pressures decrease performance without sufficient increase in air bottle yield to be advantageous.

For starters capable of outputs in the ½ to ¾ million ft-lb of kinetic energy, a combustion starter will require about 0.65 lb/sec of air while an air turbine starter will require about 1.85 lb/sec. A solid propellant liquid or liquid monopropellant starter requires no stored air supply.

4. Fuel

The fuel for a combustion-type starter is usually JP-4; however, starters have been built which operate on a variety of petroleum base fuels. The fuel requirement is usually very small and ranges from 1/15 to 1/35 of the air used. Starters are built to operate on a lean mixture to provide a clean exhaust.

The solid propellant presently used is an ammonium nitrate propellant. These are made up into cartridges complete with a case, igniter, and propellant. At the present stage of development, these cartridges are very sensitive to temperature changes and their exhaust products exhibit large amounts of smoke.

5. Electrical

Air turbine starters require no electrical power for operation. They are usually provided with a switch which will open at the designated cutoff speed, which is normally wired to the air valve to shut off air to the starter at this speed.

A combustion starter requires electrical power for the fuel valve, air valve, relays, and ignition. The ignition is usually cut off as soon as combustion has started. A typical wiring diagram for a fuel-air combustion starter is shown on Fig. 13. The fuel and air valves require about 3 amp at 24 volts during the start cycle. The ignition requires an additional 4 amp for the first ½ sec of starter operation.

Solid propellant and liquid monopropellant electrical requirements are very dependent upon the individual design of the starter.

6. Air Bottles

Air for fuel-air combustion starters is usually stored at 3000 psi in fiberglass or steel air reservoirs. Not all of the air in a bottle may be used for starter operation since the bottle is not operated down to zero pressure. The general use factor used is 0.85. In this case, the weight of air to be stored in each bottle can be figured as:

$$W_A = \frac{w_A(t + 0.3)}{0.85Y} \qquad (6)$$

The diameter of a sphere in fiberglass to contain a given amount of air can be approximated with the following equation:

$$D_B = 6.4\sqrt[3]{W_A} + 0.25 \qquad (7)$$

The weight of a fiberglass bottle can be approximated by:

$$W_B = 2.1 W_A \qquad (8)$$

7. Air Compressors

Air compressors for airborne service cannot compress air at the rate required by the starter; therefore, it must be stored for subsequent use. These

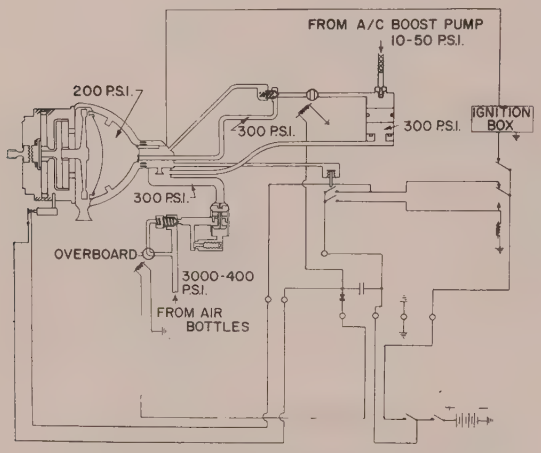


Fig. 13 — Typical fuel-air combustion starter flow and electrical schematic diagram

Table 2 — Information Needed for Starter Specification

1. Aircraft designation
2. Complete engine designation
3. Polar moment of inertia of engine
4. Polar moment of inertia of rotating accessories
5. Steady state torque versus speed curve of engine
6. Steady state torque versus speed characteristics of accessories during start
7. Idle speed of engine
8. Maximum speed of engine
9. Starter pad type
10. Direction of rotation of starter pad
11. Gear ratio, starter pad to engine rotor
12. Pad strength (continuous and static)
13. Duty cycle required of starter
14. Desired starting time
15. Allowable envelope for starter—drawing preferable
16. Approximate line distance between major components
17. Bleed air and duct availability—pressure, temperature, and flow
18. Special requirements of aircraft

Note: Complete reference must be given on torque, speed, and MOI values.

compressors usually compress 2–6 cfm (9.1–27.3 lb/hr). Air compressors can be driven by electric motors, hydraulic motors, direct engine drive, and the like, the type of drive usually being determined based upon aircraft design and power availability.

APPENDIX II

Information Needed by Starter Design Engineer

The preceding appendix covered the general equations and factors used in the design of a starting system for a turbojet or turboprop engine. The equations presented can, however, give only a general indication of the requirements and performance of the equipment. These must be modified by the starter designer to take into account other variables required by the design of a specific unit.

The starter designer is also in a position to make fairly drastic changes to the configuration of a starter in order to meet specific installation requirements. Close coordination between the starter manufacturer, engine manufacturer, and the airframe manufacturer is required in order to provide a system completely compatible with the aircraft.

The items listed in Table 2 are necessary to allow the starter designer to make a preliminary proposal on a starter drive unit. If a complete starting system is required, additional information will be needed.

Symbols

C_1 = Nozzle velocity, fps
D = Diameter, in.
g = Gravity (32.16 fps^2)
I = Moment of inertia, lb-ft^2
N = Speed, rpm
R = Turbine pitch line radius, ft
R_g = Gear ratio, starter turbine/engine pad
T = Torque, lb-ft
t = Time, sec
u = Turbine pitch line velocity, fps
W = Weight, lb (for 1 start)
w = Weight flow, lb/sec
Y = Number of units
α = Nozzle angle
β_2 = Blade exit angle, deg
η = Turbine efficiency (runs about 0.75)
η_g = Gear efficiency
ψ = Blade friction factor (about 0.85)
ω_1 = Relative blade entrance velocity, fps

Subscripts

A = Air
B = Air bottle
e = Engine
net = Summation
O = Output
S = Starter
T = Turbine
w = Windage

DISCUSSION

What Do We Want In a Starting System?

— R. C. McGuire
Eastern Air Lines, Inc.

THE FEATURE of this paper is its completeness. The authors have obviously explored and described every known or contemplated jet engine starting system. Their efforts, however, have also disclosed that no jet starter provides the flexibility, adaptability, and economy of the piston engine electric starter. Whereas the old electric starter could be used to turn the engine, motor it, and start it; the existing self-contained jet engine starters seem to be designed to do only one thing — namely, start the engine. In many instances, it has not done this too well.

It is also noteworthy that existing self-contained systems are noisy, heavy, and expensive to buy as well as overhaul. It is our understanding that a fuel-air combustor system for a 4-engine transport can weigh as much as 150 lb per engine and cost about $5000 per engine. Even the most elementary air turbine starter is more than double the cost of piston engine electric starter. We have also heard that the procedures and spares required to overhaul a combustion starter resemble those for a small turbine engine.

This is not the best testimonial for a modern engine starter system. Moreover, the number of pneumatic systems developed for starting jet engines also implies a need for better definition. In other words, are we seeking self-containment or self-sufficiency?

It is my opinion that we are seeking self-sufficiency because self-containment as presently defined leaves much to be desired. The self-contained combustion or combustor starter can run out of air storage; the monopropellant starter can run out of fuel, the cartridge starter out of ammunition, and the hydraulic starter out of accumulator or electric power. If the engine is difficult to start, or if a hot start occurs, the energy level of existing self-contained systems is generally inadequate.

Hence, we need more sufficiency, along with self-containment.

True self-sufficiency can only be acquired with an airborne gas turbine compressor (GTC) because if the GTC runs, it will provide all the air needed to start or motor one or more jet engines. It is obviously best-suited for a bomber or a transport because it can furnish electrical power for ground lighting, radio, and air conditioning when not being used to start engines. It can also be adapted to motor engines for line maintenance inspection.

It will probably cost $25,000 to install and weigh over 600 lb, but part of this expense and weight will be offset by replaced systems. It can replace a ground power unit which costs over $15,000, a ground starter which costs over $20,000, and possibly an air-conditioning truck which is another $20,000 machine.

Needless to say, the airborne GTC is not suitable for a fighter airplane but the need for sufficiency in a fighter is none the less. This is substantiated by the greater interest bing shown in combination starters. The combination starter does not provide self-sufficiency but it does provide sufficiency for all ground and starting needs when supported by a gas generator ground unit. For normal daily operations, the low torque portion of the starter and ground unit provide excellent support; for a scramble or airline gate departure, the high torque section of the starter can be used.

In closing, we recommend that the industry henceforth develop one jet engine starting media to maximum sufficiency, reliability, and economy. The existing systems just don't meet all requirements. If we ever decide to develop a new system we should give the hydraulic starter another look.

possibility of "Spare Tire by use of Dual

Walter J. Lee
The Goodyear Tire & Rubber Co.

THE PRINCIPAL REASON for wanting to eliminate the fifth wheel on a passenger car is to make that space available for other purposes. There are other reasons, such as greater freedom in styling the rear deck.

Before this last spare part can be eliminated, the *need* for it must be eliminated — at least for all practical purposes. To oversimplify, let me say that present tires must be made still safer, still more dependable, more puncture proof, more blowout proof, and more reliable in every way before the public will let go of the present spare tire.

We think all this can be done and we would like to tell you how we are going about it, what approaches we are taking, what we are doing to get ready for the time when four tires may be standard equipment on automobiles.

Years ago, as some of us will remember, spare tires were decreased from three to two and then finally down to one. This elimination took place as rapidly as the art of making tires progressed. As tires became safer, spare tires disappeared.

Fifty years ago three spare tires were standard practice and frequently all three were used on a Sunday afternoon. Thirty years ago two spares were still in common practice, but not used often enough to justify both of them. Today, the fifth wheel in the trunk is being used less and less as an emergency spare. It is used mainly in rotation with the other four.

Psychologically, however, it is as necessary as ever. The peaceful assurance of knowing there *is* a spare tire in the trunk is as important today as it ever was — almost as important as liability insurance. It is one thing to eliminate spare tires from three to one, but quite different matter to eliminate that one down to zero. The *possibility* of this total elimination of the spare tire is the subject of this paper.

The first question usually put to us in a discussion of this kind has already been answered. As good as they are, the four working tires you now have on the ground are still not sufficiently reliable to eliminate the need for the fifth tire in the trunk.

The second most natural question is, "How close are we to the time when four tires *will* be sufficiently reliable so that the general public will accept the automobile without a spare?" I don't believe anybody can answer that question to your complete satisfaction today. Certainly, I can't. It won't happen in 1960 and probably not in 1961. In 1962? Possibly to a limited extent and then probably on an optional basis. At least at the start.

Just like the passing of the hand crank and the air pump, and the tool kit, the fifth tire will be dropped when the majority of new car buyers feel it is no longer necessary for their peace of mind. Now then just what are the immediate chances of four tires developing sufficient confidence in the minds of the driving public so that a fifth tire will not be demanded by the majority. In our opinion

Elimination"

Compartment Tires

SPARE TIRE elimination is a problem that has defied tire engineers for years. Because the space it occupies is wanted for other purposes, increased efforts are being applied to the search for a solution.

This paper describes one approach — the dual compartment tire. This is made up of two tires, one inside the other, both of which are inflated at all times. If the outer tire is punctured, the inner tire supports the car until it reaches a service station.

According to the author, tests indicate that this solution holds promise of eliminating the spare — if the cost can be lowered.*

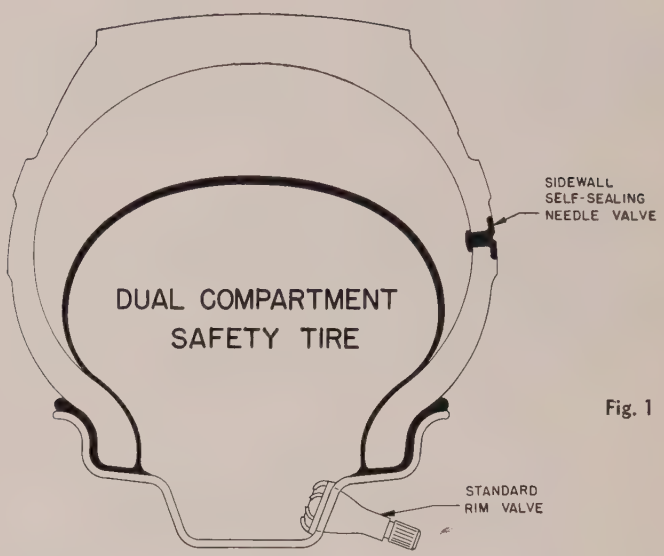


Fig. 1

the dual compartment principle of tire design has the best chance of accomplishing this goal. (See Fig. 1.)

The outer tire is of standard shape and dimensions. The inner tire, or safety shield, is made of two plies of nylon fabric reinforced through the riding area with two plies of wire for puncture protection.

The inner tire is inflated through the rim valve as per present practice. The outer tire is inflated through a needle valve built into the sidewall. This operation requires a special adapter for the standard service station air chuck.

Goodyear has had this dual compartment type of safety tire in regular production under the name of "Captive-Air" for more than two years.

As you can see, if one compartment is punctured or cut, or fails to hold air for any reason, the other compartment takes over the load and keeps the car going without a road delay. If, however, the road hazard is big enough to slash through both compartments, the dual tire principle has failed. Thus, the success of this new concept of a safety tire depends on a low incidence of failure to both compartments.

How often will both chambers be ruptured at the same time? If this happens at an average of once per 100,000 car miles, we would not consider it acceptable. That is too often. If the incidence of failure is once in 200,000 car miles, we would feel

* Paper presented at SAE Summer Meeting, Atlantic City, June 17, 1959.

we are getting close to the realm of acceptability. So like every other part of the automobile, the factor of safety spells the success or failure of the dual compartment tire.

Now a word about where we stand in producing this safety tire that we think may someday eliminate the need of a spare tire. We have approximately 900,000 of these dual compartment tires in commercial use on automobiles today. These tires are being sold in the replacement market under the name of "Captive-Air" for maximum insurance against road-side delays, maximum safety, maximum dependability. The car owners who have bought these dual compartment tires already have spare tires and in most cases they are not removed. However, we are hearing of more and more cases where the spares are being removed to make way for additional luggage. We think the percentage of these cases will increase as more and more confidence is developed by the trouble-free performance of these safety tires.

In addition, 166 taxicabs are using the safety tires — without spares. These cabs are running in Erie, Pa., San Francisco, Los Angeles, New Orleans, and Birmingham, Alabama, — a good cross-section of the country. In general, the tire performance is well above the standard tubeless tire, but we *have had* some failures.

I will present the current results of this broad scale road test and let you form your own opinion on whether or not the dual compartment tires have a chance of someday eliminating the need for the spare tire. Tire companies, as you probably know, are paid by taxi companies according to the number of trouble-free miles the tires deliver. Conversely, we are penalized by the number of road side delays caused by the tires. Table 1 shows the latest results we have on the keep-going ability of our dual compartment tires. This table shows that if dual compartment tires had not been used, these 166 cabs would have been down 1776 times because of flat tires during these 19 months. Actually however, 1707 times the secondary inner tire (the safety chamber) was not punctured — did not go flat — and the cabs were still able to operate under their own steam.

The remaining 69 times, the cabs *were* down. The dual compartment principle failed to work, both compartments were ruptured. This data certainly indicates a tremendous safety advantage in favor of the dual compartment principle over the present standard tubeless tire.

Another comparison is based on over 1000 cabs in 11 cities across the country and it goes back five years to the time when tube-type tires were in common use. (See Table 2.)

There is no question that these figures point out increased safety of the dual compartment tires over standard tires.

We believe the dual compartment principle of tire construction offers excellent protection against punctures and blowouts. Road side delays due to tire failures are reduced to a minimum.

We don't feel the present Captive-Air tire is as good as we can make it. We have a group of engineers, both chemical and mechanical, who will continue developing the dual compartment principle for still greater tire dependability.

Table 3 shows data that has a bearing on our subject. We feel that if dual compartment tires were in common use, AAA calls for road service on flat tires would drop down to about the second asterisk.

Our *aim* is to be able to sell four dual compartment tires for original equipment at a price not to exceed the price of five present tires plus an average of about $11.00, depending on size.

The $11.00 is our own estimate of what the extra wheel would cost plus the jack, and plus any other expense the car manufacturer would be put to to accommodate the spare tire. This is our aim, but we have not reached this goal yet. However, I think it is in sight.

In conclusion, let me say that notwithstanding the rather positive statements I have made in support of the dual compartment principle for maximum tire safety, we have not closed the door on other ways of eliminating the present spare tire. We are still working on other approaches, but so far they all require a road side tire change — the difficult inconvenience we would like to avoid. The dual compartment principle keeps the stand-by tire in place at all times, and that is why we favor this type of tire to eliminate the need of the present spare tire someday.

Table 1 — Dual Compartment Dependability in Taxicab Service

1. Number of cabs in test	166
2. Length of test, months	19
3. Number of tire miles	42,300,000
4. Total number of flats	1,776
5. Total number of times outer compartment *only* went flat	1,707
6. Total number of times both compartments went flat	69

Table 2 — Dual Compartment Tires Safety Comparison

	Average Tire Failures per Cab per Year
Tube Type	38
Tubeless	11
Dual Compartment	⅛

Note: The overall average for taxicabs is approximately 50,000 miles/year.

Table 3 — American Automobile Association Emergency Road Service Calls — 1958

	Percentage
Battery and Electrical System	22.1
Tires	21.1*
Ignition	13.8
Tow and Wreck	10.0
Stuck	5.0
Carburetor	3.6
Starter	4.3
Out of Gas	3.0
Gas Line	2.7
Brakes	2.1*
Lock and Key	1.6
Lights	1.1
All Others	9.6
	100.0

(Total number of 1958 calls — 60,497,000.)

the expanding

Polymer Horizon

J. H. Du Bois

Tech-Art Plastics Co.

THE DEVELOPMENT of new polymers offering new properties and new combinations of desirable characteristics, coupled with advances in manufacturing techniques, has expanded the plastics horizon. This paper describes some of these new materials and a few of their many possible applications in the automotive industry.

The author emphasizes that greater use of plastics in the automotive field depends to a great degree on the imagination and ability applied in creating new products. Design features most overcome the fundamental limitations of the new materials. The basic weaknesses of plastics are listed.

Production techniques will affect the future expansion of the industry. Three methods show particular promise: blow molding, fluidized polymer deposition, and potting compounds.*

* Paper presented at SAE Summer Meeting, Atlantic City, June 19, 1959.

IT IS OFTEN difficult to tell when a polymer is new. Plant samples and literature on the new products are distributed several years in advance of the materials' availability. These so-called "new materials" are usually "old hat" in the Research and Development sections when they are first put into commercial use. We include many such methods and materials in this paper in order to present the advanced practice in the manufacture and use of the polymers in this period of accelerated materials' development.

It is to be expected that the material makers will recite the virtues of their products and minimize their weaknesses. Designers are often blinded by the attractive features and expected economy. They fail to recognize certain of the fundamentals which must be considered in connection with any new product design. These fundamental limitations of the plastics must be met by design features. The following are well-known sources of failure with the organic plastics that must be recognized in every new product development and considered carefully in the valuation of a new polymer.

1. Lack of absolute age and dimensional stability.
2. High and variable thermal expansion.
3. Effects of moisture absorption.
4. Loss of plasticizer.
5. Weatherability.
6. Batch-to-batch material variation.
7. Flammability.
8. Low scratch resistance.
9. Low thermal endurance.
10. Effects of flame, electrical arc, and nuclear radiation.

Certain manufacturing procedures are of special interest in connection with the expanded use of the polymers since much of the future volume will depend on the new production methods.

Blow-Molding Developments

It is expected that the improved machinery now being built will permit the blown plastics to compete

with glass and metal containers in the near future (Fig. 1). Improved materials and blowing machinery have indicated several expected automotive uses for blow-molded plastics. Typical items in this field include the following:

Windshield visors and arm rests can be blow molded of polyethylene or polypropylene that will have good safety features, desirable tactile surfaces, light weight, and reasonable cost.

Increased use of plastic bottles for windshield wash liquids will come with cost reductions and more functional designs. For example, high-density polyethylene gallon bottles now cost 25½¢ compared with glass at 28–30¢. Accordion-shaped bottles can be caused to discharge by a simple mechanical linkage.

Many studies are being made to make advantageous use of plastics in the brake and shock absorber systems. A high-density polyethylene or a nylon blow-molded vacuum cylinder are very real possibilities.

Several improved materials and treating processes have indicated the future potential of blow-molded containers for motor fuels. While considerable work lies ahead in solving the permeability problems, the potential is very real. Costs will be attractive and shapes may be achieved that will make maximum use of irregular space areas. (Fig 2 is an example of a bottle possible by the method.)

Bellows type "bottles" will prove to be economical for foot pedal windshield activators and for the accelerator and brake seal (Fig. 3).

Potential use of blow-molded plastics in the automotive field is aided by low mold costs, fast tooling, and production technique flexibility which permits design variation, color, and low product cost.

Fluidized Polymer Deposition

The fluidized bed technique consists of immersing the product, after heating above the plastic's melting point, in a bed of fluidized plastics resin. After its removal, the residual heat in the product serves to fuse the adhering plastics into a continuous film.

Polymer Processes, Inc. who developed this procedure have demonstrated its feasibility with polyethylene, polystyrene, nylon, acrylic, polytetrafluoroethylene, and acetal resin.

Potting Compounds

A variety of potting and casting compounds have been developed that simplify many construction and assembly programs. They offer substantial help in prototype work. Included in this group are high-temperature resins, and improved impact, flexible, and lightweight foam products. New transparent and general purpose casting resins are on the market. 500 F epoxy sealing compounds, which may have adjustable flexibility, are being extensively used in electronic and engine assemblies.

It is possible that ignition coils of the future will have no molded case, but be potted in an epoxy or polyester compound to improve cooling, eliminate moisture penetration, and reduce cost. The distributor cap may become an integral cast unit, composed of the several wires having crimped-in-place terminals, permanently cast in a minimum-cost potting compound. This would end moisture penetration and be a throw-away item when replacement is indicated.

Techniques already accepted for electronic and aircraft products will introduce some of these useful polymers for many sealing purposes; their cost will be justified by simplified components and assembly.

Phenolic Materials

The phenolics continue to be basic construction materials for many nonmetallic and insulating products.

Sisal-filled phenolics offered by Durez and Rogers are gaining good acceptance as a replacement for rag stock materials and certain polyester-glass premix. This is inexpensive material that may be molded at low pressure in cast-steel molds. It is being used in auto heater housing and air-conditioning components; and may be employed in some wet applications where the conventional phenolics fail.

The glass-filled phenolics are gaining acceptance rapidly; typical applications are the familiar Durez-Chrysler oil pump gear — whose performance has been proved to be superior to that of die-cast aluminum. Automatic transmission cones produced from this material are molded directly to tolerance without additional machining. It transmits full engine power to the rear wheels of one of the heaviest passenger cars.

A modified phenolic (Fiberite 4035 natural) with glass fiber reinforcement is performing well in a number of critical applications that require dimensional stability, minimum thermal expansion differential, high strength, and heat stability. It is suited to the molding of components which require complex insert assemblies that must maintain tightness over a considerable temperature range. (See Table 1.)

A new printed circuit material, Continental Diamond — "Di Clad 2350 P," has electrical properties superior to Nema Grade X and is lower in cost. It can be fabricated by punching and dip soldering to produce reliable automotive wiring systems at reasonable cost.

A new two-stage, mineral, and flock-filled phenolic molding compound "GE 12933" exhibits improved resistance and dimensional stability, good heat resistance, and excellent electrical properties.

A nylon flock-filled two-stage phenolic molding material, "Rogers RX 725," is particularly suitable for mechanical applications such as gears, bearings, and joints. A graphite-filled variety has also been introduced.

Styrene Plastics

The styrene materials make the lowest cost thermoplastic molded products. The new Foster Grant graft polymers exhibit great strength and are being used increasingly for decorative components that require improved abuse resistance. High-impact styrene will be used with a 60% cost reduction for injection molded ventilating shrouds on all 1960 Fisher bodies. This work was formerly done by compression molding polyester premixes. It represents the first important automotive use of styrene since 1942, when it was abandoned because of its poor solvent resistance.

The styrene ABS (acrylonitrile-butadiene-styrene) polymers, "cycolac" by Marbon Chemical

Division, offer maximum toughness, high hardness, and tensile strength. Styrene ABS resins may be injection molded, formed from sheet, and fabricated by all conventional procedures. New potential uses included bezels, inner door panels, steering wheels, roof liners, and trim. Kralastic MM, a new ABS resin by U.S. Rubber, shows extremely high elastic modulus and other desirable properties that will enable it to compete with the acetal and nylon resins.

Polyethylene

The several polyethylenes have been separated by ASTM D-20 into three types:

Less than 0.910 g/cc — Lowest density
Type I 0.910–0.925 g/cc — Lower density
Type II 0.296–0.940 g/cc — Medium density
Type III 0.941–0.965 g/cc — Higher density
Above 0.965 g/cc — Highest density

A whole new series of copolymers and alloys of the polyethylenes will introduce important additional values — such as more uniform shrinkage, greater transparency, plus added resistance to environmental stress cracking.

New applications for polyethylene include vacuum-formed air-conditioning ducts, windshield defroster nozzles, windshield washer reservoirs, sun visors, heater adjustment cable and sleeves, graphite-filled linear spring interleads, plugs, antenna grommets, glove compartments, dashboards, deck trays, gas tank caps, formed seats or components, and air intake tubes for filters. Polyethylene has been found to make a good door water shield. Barbed shank, self-attaching polyethylene fasteners are being used increasingly for attaching dash liners and to fill holes during shipping.

In using polyethylene, emphasis must be placed on proper test procedures and the establishment of suitable molding techniques. Low-temperature impact strength and stress cracking resistance are obtained by proper molding techniques. The plastograph has proved to be an excellent means of

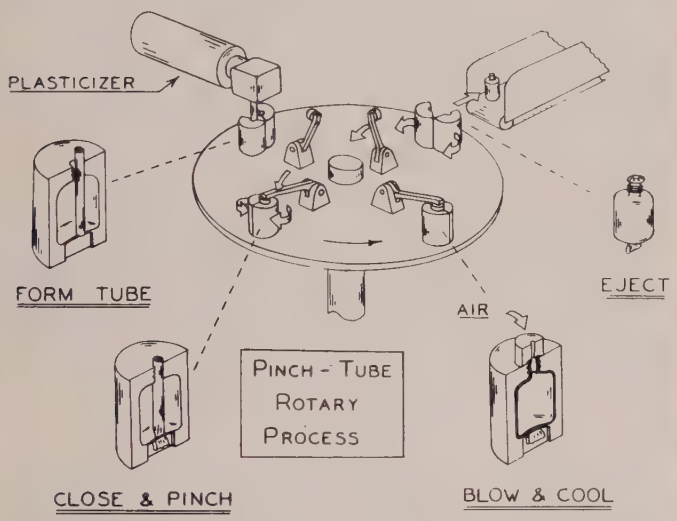


Fig. 1 — Pinch tube rotary process developed by Plax Corp. for blow molding hollow objects. The plasticizer extrudes tube of plastic material. As die closes, bottom end of tube is pinched and air is injected at top to "blow" shape. Cooling takes place as mold moves around table. The project is ejected from mold by air blast in last stage

Fig. 2 — Blow-molded "Jerican" illustrates functional reinforcement and integral attachment features achieved in single blow-molding operation

Fig. 3 — Blow-molded bellows bottle dispenses gases or liquids by compression. Many plastics offer long flex life

Table 1 — Thermal Expansion Coefficients for Plastics and Insert Metals

	10^{-6}/deg C		10^{-6}/deg C
Electrical glass	0.8–13.0	Monel	14.0
Porcelain	5.0	Type 18-8 stainless steel	16.0
Silicone-glass laminate	5.0–10.0	Copper	17.7
Steatite	8.6–10.5	Nickel-silver	18.0
Platinum	8.8	Red brass	18.8
Illium R	9.2	Mica-filled phenolic	19.0
Glass-bonded mica	10.0	Silver	19.1
Polyester-glass	10–30	Melamine-asbestos filled	20–45
Type 446 stainless steel	10.5	General purpose phenolic	30–45
Ceramoplastic	11.2	Alkyd-mineral filled	35.0
Inconel	11.5	Polymonochlorotrifluoroethylene	45.0
Type 430 stainless steel	11.8	Polystyrene	60–80
SAE 1010 steel	12.1	Polytetrafluoroethylene	99.0
Nickel	13.0	Nylon	100–150

forecasting the processability of polyethylene and the other plastics.

A three-dimensional, structural fabric "Trilok" is produced by U. S. Rubber Co. using the differential shrinkage between polyethylene and vinylidene chloride or nylon fibres in the weave. It shows promise of increased use for lightweight structural sections, filtration purposes, and for self-ventilating seat cushions.

Polypropylene

Polypropylene is a heat-resistant, hard, glossy, chemically inert, and lightweight (0.90 specific gravity), moldable and extrudable polymer. It will compete with polyethylene, the cellulosics, vinyls, and styrene. In view of its good heat resistance, it will compete with nylon in certain previously unchallenged fields.

Polypropylene film has greater tensile strength than any of the various density polyethylenes. Its impact and tear strength is in the range of the lower density polyethylenes. Its stiffness is equivalent to or greater than high-density polyethylene. It is heat sealable and its low-temperature flexibility point — 80 F — is within the limit of automotive use. Its gloss and clarity are exceptionally good and it is boilable.

The markets now being developed include automotive and electrical components, automotive textile fibers, steering wheels, pipe, distributor caps, dome lights lens and bezels, valves, fittings, fans and housing, films, and the like, which take advantage of its fine gloss, improved mar resistance, light weight, and freedom from stress cracking. Thermo-formed sheet polypropylene materials may be used for luggage, housing, duct work, and glove compartments.

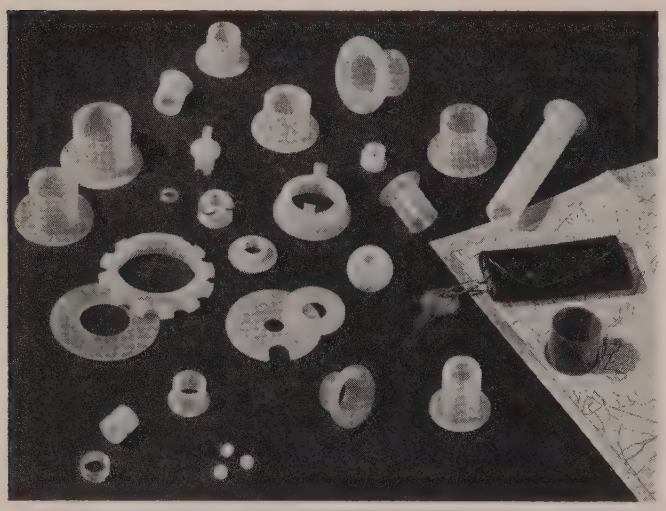


Fig. 4 — Bushings for brake and clutch pedals, window cranks, steering columns, door hinges, and suspension systems are standard uses for nylon resin. Car makers' requirements are rugged: Window bushings must withstand 50,000 cycles; steering column bushings, 1,000,000 revolutions of wheel; and clutch bushings must still be serviceable after 400,000 cycles. Nylon is tough in sense of being resilient and is properly used where ability to yield under blow is important

Acrylic Resins

Polymers of esters of acrylic and methacrylic acids are widely used for the coating of metal products. These solution polymers dry to produce weather-resisting clear, hard, tough protective coatings. They may be pigmented to produce very attractive colors.

The weatherability of the acrylic resins accounts for many of their automotive applications.

A new acrylic-type polymer which can be injection or extrusion molded is of special automotive interest because of its high thermal endurance and good combination of stiffness and toughness. These materials compare favorably with the conventional acrylic resins but gain 50 F in heat distortion temperature and may be immersed in boiling water indefinitely without loss of clarity. Products of this type are identified as PL-11 and 12 by J. T. Baker Chemical Co. and as "Zerlon" by Dow Chemical Co.

A modified acrylic resin for injection and extrusion is offered by Rohm & Haas under the name of "Implex." It is characterized by improved toughness without loss of colorability and gloss; Implex contains no plasticizer.

Polyester Products

"Mylar" polyester film is a transparent, flexible du Pont plastic material having broad and useful properties. Its tensile strength exceeds 20,000 psi and it is insensitive to moisture, chemicals, and solvents. Its dielectric strength is above that of other flexible plastics and it will operate between

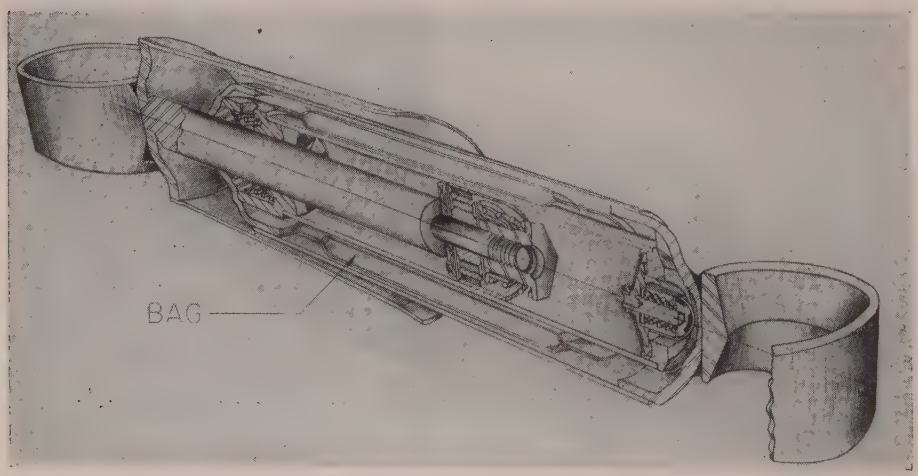


Fig. 5 — Nylon 6 is used for Delco Products "Pliacell," pliable cell enveloping reservoir gas. Freon 13 gas, sealed in this nylon bag, is alternately compressed and expanded within bag by motion of shock absorber piston

– 140 and + 300 F. Since there is no plasticizer, there will be no age embrittlement. Adhesives will bond it to itself and other materials. Metallized, polyester film is used to obtain attractive metallic effects in upholstery materials and interior trim. Tough and flexible nameplates, metallized and printed, are inexpensive, attractive, and serviceable. Slot wedges of thin polyester film permit maximum copper in minimum generator and motor slots with adequate insulation.

Polyesters derived from isophthalic acid by Oronite Chemical Co. exhibit higher softening points than Mylar films which will expand their fields of application.

Polyester resin premix materials such as Reichold's "Polylite" are gaining favor for door panels, glove compartments, truck cabs, doors, and bodies. It is expected that the newly developed, very tough materials will serve as headliners and door liners — with the decorative fabric molded into the resin; they will be adhesive bonded to the doors and roof. Reinforced seat forms will be used, filled with shaped, flexible polyurethane foam, and covered with adhesive-bonded, textured plastic covers. This combination will provide comfortable, springless seat assemblies.

While the number of polyester resin applications is not large, considerable volume of premix material has been utilized for heater ducts and for miscellaneous engine and body parts using sisal fibers and glass fillers. Typical of the increased applications are door panels, glove compartments, and paneling.

A glass roving-filled diallyl phthalate product has high impact strength and may give improved servive in ignition and electronic applications requiring stability of properties under high humidity conditions.

Design integration permits the substitution of polyester products for steel with economy. The development of self-extinguishing polyester resins has improved their market position.

Nylon Plastics

The excellent performance of nylon products has expanded their automotive usage to an average quantity of 42 nylon parts per car (Fig. 4). Nylon and "Delrin" will be competitive in some applications. "Delrin" is tough in the sense of being strong and rigid and is properly indicated where high strength and resistance to bending is required. Nylon is tough in the sense of being resilient and is properly used where the ability to yield under a blow is important.

Nylon 6 materials are widely available and their resilience offers some advantage in certain classes of work. A new hydrolysis-resistant nylon resin, Zytel FE2281, is suitable for use in the cooling system; a typical application is an anti "gurgling" device mounted in the heater base inlet assembly.

Nylon tubing is increasingly used in the air-suspension and vacuum spark line. Nearly completed test data indicate far more extensive use of nylon tubing.

Glass-filled nylon such as "Fiberfil" by Fiberfil, Inc. has increased tensile strength amounting to 50%, inpact increased by 122%, heat distortion temperature at 264 psi is increased 230%. Deformation under load is cut to one-third and the linear thermal expansion is reduced by 30%. Glass-filled nylon has reopened the bearing cage application to the thermoplastics. Numerous new small part applications take advantage of the glass reinforcement in nylon.

A new Delco shock absorber, makes use of a freon-filled Nylon 6 bag which serves as the cushioning medium (Fig. 5). As fluid is displaced in accordance with movements of the piston, the gas is alternately compressed and expanded within the bag, eliminating mixing of gas and fluid. Fosta-Nylon 6 film is used for the bag which has performed nicely over a two-year period.

Sintered Nylon

Sintered nylon products are gaining ground in the mechanical product field; they are competing with sintered bronze, steel, and many of the porous metal combinations. These products are combinations of nylon, graphite, molybdenum disulphide, or are made porous and used with an oil impregnation. These products are produced by the Nylasint process of National Polymer Products, Inc., which is based on cold pressing and sintering techniques similar to powder metallurgy. New applications include bearings, cams, gears, seal rings, thrust washers, rollers, sliding shoes, piston rings, pump vanes, and the like. On a typical 50,000 (50M)-per-month item, it was found that tooling costs for a sintered nylon cam are less than one-third the cost of tooling for an injection-molded nylon cam.

Delrin-Acetal Resin

"Delrin" polyoxymethylene is expected to have a very large future as a result of its stability against aging, chemical combinations, and depolymerization. "Delrin" acetal resin is a du Pont product with high strength that will have numerous applications in the mechanical components field where it will replace many of the metals. It offers high strength, stiffness, colorability, toughness, low static and low dynamic friction, solvent and abrasion resistance, low permeability, and good dimensional stability. In strength, it has both high tensile strength and high flexural modulus combined with resilience and toughness. Electrical properties will be satisfactory for most applications. Resistance to hydrocarbons is good throughout the range. Acetal resins are slow burning.

For automotive purposes, many successful tests have been run using "Delrin" parts under the hood at operating temperatures up to 250 F and higher. Widespread use of the acetal resins will come as a replacement for zinc and aluminum die castings for functional and decorative service. Serious programs are under way for its use in instrument cluster housings, interior hardware, carburetor and fuel pump components, windshield wiper-washer parts, and for bearings, cams, and slides.

Polycarbonate Resins

The polycarbonate resins are becoming increasingly available and have satisfied a number of unusual programs. "GE-Lexan"' is a typical product with outstanding creep resistance; the creep of nylon at *room temperature* is no better than that of "Lexan" at *100 C* under similar loading conditions. "Lexan" is more creep resistant than "Delrin," a

significant factor is that "Lexan" reaches its point of best dimensional stability in less time than is required for "Delrin" and nylon to attain their state of best dimensional stability. A clear polycarbonate is produced by Mobay Chemical Co. and is suitable for hot lamp lenses, and the like. Dimensional stability, improved impact strength, and heat resistance up to 300 F are important values obtained through its use. Typical applications include gears, bottles, electro-mechanical components, coil forms, cams, lenses, and the like. These resins are self-extinguishing.

Coil forms of polycarbonate will withstand temperatures of 280 F without deformation or oxidation. High precision mechanical components may be molded advantageously of polycarbonates. All of the conventional automotive finishes and metallizing may be applied.

Epoxy Resins

The epoxy resins have already gained wide acceptance for potting, adhesive, and bonding applications. Numerous epoxy-resin "welded" metal joints have proved their high bond strength. Near hermetic seals are achieved by vacuum impregnation with epoxy resins. New uses will include instrument fastenings, small components, and trim. These bonds are vibration resistant and more secure than bolts or solder.

A tough, epoxy based enamel with very high bond strength to molded phenolic compounds is offered by Rexton finishes. It is available in a glossy or in "zero gloss" for nonreflective coatings needed in cameras and for panels. Adhesion is good, even to very high gloss "as-molded" phenolic surfaces.

A new development in automotive practice is the use of cast epoxy core boxes "Epocast 11B" made by Furane Plastics, Inc. It is highly resistant to the abrasive action of sand blowers and is often cast about a foundry pattern and confined in the metal flask. No machining is needed after casting. Another epoxy casting and trowelling compound, "Masterlite 320," is a close match for mahogany and is being extensively used in some automotive pattern shops as a superior material. It may be drilled, nailed, and fabricated by wood-working tools.

A crystal clear epoxy casting resin by Marblette called "Maraglas" (Fig. 6) is excellent for model and prototype work, being particularly valuable for prototype lenses and other acrylic products.

The tooling potential with the epoxy resins is the subject for a major paper and is worthy of special study by those who have this problem.

Several new epoxy molding compounds are on the market with most desirable stability properties.

Epoxy-Bonded Inserts

Use of the epoxy adhesives offers large possibilities for thermosetting product savings. By proper hole and insert design, inserts may be epoxy-bonded in place after molding, with great holding power and freedom from leakage. One such job which was produced with molded-in metal inserts in a 6-cavity semi-automatic mold was changed over to a fully automatic molding process, with the inserts cemented in place after molding. Costs were reduced from 18¢ to 8½¢ by the resulting mechanization of the process. Other cost reductions have been achieved by post-molding cementing operations to facilitate fully automatic molding

Silicone Plastics

One interesting material Silastic RTV 502 is a fluid, quick setting, room temperature vulcanizing silicone rubber. It remains rubbery from – 70 to – 500 F, has excellent electrical insulation, resists ozone, weathering and moisture, and absorbs mechanical shock. It is used for sealing, caulking, and potting and for making impressions of objects. Such impressions can then be used as a mold to duplicate the original.

Several silicone sprays offer a high degree of water repellency to protect electrical components. Silicone rubber insulation is most excellent for high capacity electrical generators — it has high thermal conductivity and good electrical properties from – 130 F up to – 500 F. Molded silicone rubber components can operate very close to engines where other plastics fail.

Teflon 100X FEP Fluorocarbon and Teflon TFE Fluorocarbon Resins

"Teflon 100X" is one of the more important of the new technical materials; it provides a perfluorocarbon resin that can be molded and extruded with conventional equipment.

Teflon 100X has the desirable properties generally associated with "Teflon" but with a reduction in operating temperature limit to approximately 400 F. It can be combined with various inorganic fillers to gain rigidity, dimensional control plus a minimum loss of its most desirable electrical properties. For molding purposes, it may be pre-heated at 750 F and cooled in molds at 400 F. Injection or transfer-type molds can be used.

Automotive applications will include printed wire, seals, gaskets and bearings, engine harness jackets, hydraulic and fuel lines, corrosion resistant coatings, and all anti-friction surfaces.

The FEP and TFE fluorocarbon resins exhibit most desirable properties for automotive applications:

1. Stability of properties at elevated temperature.
2. Inertness to solvents, chemicals, and corrosive gases.
3. Very low coefficient of friction; static friction and dynamic friction are the same.
4. Excellent resistance to electrical arc and flashover.
5. Unlimited resistance to weathering and aging.
6. These resins are noninflammable.

A most interesting new material is Duroid 5650 manufactured by Rogers Corp. Ceramic fibres, 50% silica and 50% alumina, are coated with polytetrafluoroethylene by the beater addition process and then molded into a homogeneous sheet. The ceramic fibres minimize cold flow, improve thermal conductivity, and permit this product's use in high flange pressure seals and gaskets and for highly loaded bearings. Other varieties are porous and contain oil or graphite. New applications may be found in the oil and hydraulic circuits of automobiles.

A glass micro-fiber reinforced Teflon, "Duroid 5613" exhibits good wear resistance and minimum cold flow. Sleeve or thrust washer type bearings operate nicely at temperatures up to 550 F under corrosive conditions where lubricants can not be used.

A coating which consists of a mixture of phenolic and TFE resin has been introduced by Acheson Colloids Co. It may be applied by brush or spray to plastics and other heat sensitive materials. On heat curing, the coating is firmly bonded and has the low-friction and flash-resisting properties of the TFE resins.

Foamed Plastics

The foam-in-place plastics are produced from many basic resins. Urethane foams are gaining in acceptance because of their favorable cost factors. They are cast in flexible, rigid, and semi-rigid sections.

A typical foam can be made to stretch 700% of its original length, withstand 150-psi pressure, and perform nicely from – 30 F to + 250 F. Properties can be modified as needed to meet a given requirement. Cast-in-place foams have provided the essential thermal insulation for trucks hauling frozen foods.

New applications will include foam-in-place weather stripping and sound insulation on the fire wall and under the car body. Recent tests on airless automobile tires filled with a special urethane foam have passed some of the most severely punishing tests and it is hoped that continuing investigation of the foams will end the blowout hazard.

An expanded polystyrene bead product, offered by Dow Chemical Co. under the trade mark name "Pelaspan" and "Dylite" by Koppers Co., is finding some most interesting markets. This fills the need for a foam-in-place material to supplement the styrene foam blocks previously available. These beads contain an expanding agent and when confined in a retaining mold and heated, they expand and knit to form a cellular rigid foam product. Very substantial markets are expected in refrigerated truck bodies. Koppers has developed techniques for converting molded Dylite into a soft, resilient foam material which can be used in automobiles for floor carpet underlay, dashboard crash padding, head liners, arm rests, door paneling, and the like.

The rigid polyurethanes are being used advantageously for weight reduction in structural applications with and without Freon 11 as a foaming agent for thermal insulation and in crash pad applications. The new one-shot process by Mobay Chemical Co. and du Pont has helped this market.

Vinyl Plastics

A new polyvinyl Fluoride film, type R, is being produced experimentally by du Pont. A feature of this transparent film is its exceptional weathering stability, chemical resistance, and mechanical strength. It can be vacuum metallized, thermo formed, and laminated.

U. S. Steel has introduced a vinyl plastisol coated metal sheet which may be handled like steel. It may be drawn 30% with no ill effects and with no under film corrosion. A 5½ in. diameter test cup was drawn to a 3½ in. depth without damage to the vinyl. Colors and attractive textures are possible. It is stable up to 160 F continuous exposure or at 212 F for two days. It will find markets in fabricated steel parts which require a corrosion and highly wear-resistant coating. Welding procedures have been developed that permit fusion of the steel without damage to the vinyl.

Fig. 6 — Almost limitless range of applications for which transparency is desired is claimed for new "Maraglas" epoxy resin which can be cast easily in any size or shape, or used as laminating material or adhesive

Vinylidene chloride (Saran) fabrics are making an excellent gas fuel filter. They are bonded with dielectric heating.

Pyromellitic Dianhydride by du Pont, when used as a crosslinking agent has produced a new family of plastisols which can be made rigid or flexible. These offer certain processing advantages that will result in their increased use.

Chlorinated Polyether

A high molecular weight-chlorinated polyether thermoplastic, under the trade mark name "Penton," is being offered by Hercules. Penton exhibits a high degree of abrasion resistance, excellent dimensional stability which is not affected by moisture pickup and an ease of fabrication that is enabling it to compete with many metals in the corrosion equipment fields. It may be applied by several processes as a coating material.

The markets under development include valves, pipe, precision gears, carburetor and generator parts, protective coatings, and tanks.

It is obvious that future automotive use for the plastics will depend to a reasonable extent on the new polymers and to a greater extent upon the imagination and ability of both industries to create new products by design integration. We must conceive and demonstrate how to combine the functions of several parts into a simpler, more functional, and less expensive product combining metals and nonmetals by the best fabrication techniques which create new values. Experience has taught us that the automotive industry will not buy a better product if the present material is adequate and costs less. The plastics industry must demonstrate its new savings to gain a greater share of the automotive market.

Needs of Automotive Industry

While some of the following items are partially developed, continuing work is needed to gain wide industry acceptance:

1. An organic body solder with better cold impact

properties than the presently offered epoxies.

2. An adhesive that is capable of bonding emblems and decorative moldings to car bodies.

3. A plastic capable of replacing steel in the manufacture of car bodies. Glass-reinforced polyesters have excellent properties but the required hand lay-up method of manufacture is not suitable for large volume production.

4. An adhesive that would no longer require riveting and welding operations.

5. An abrasion resistant and shatter-proof plastic as a substitute for glass.

6. A good metal plating process for the phenolics.

7. Advanced studies to find the most effective methods for predicting the processability of the plastics and to insure product uniformity.

DISCUSSION

Review of Plastics Industry Show Rapid Progress

— W. O. Bracken
Hercules Powder Co.

THIS PAPER has very effectively attained its goal of depicting new areas for commercial exploration of polymers. In a discussion of this presentation, perhaps the most valid comments would be those arising on the first re-reading and then let the train of thought started by the author roll into new areas (new, at least for the discusser) of possible automotive usage.

The comments regarding the general lack of comprehensive engineering data for plastics state a problem recognized by the industry as a whole and a problem being worked on by the materials, manufacturers, processors, and end users. If there is this cognizance of a difficulty, why has it not been resolved, at least to a greater degree? The answer, in great part I believe, is due to the dynamic nature of the field. In a sense, not only will the whole family of plastics refuse to stop its expansion, but of the very individuals in the family, each insists on markedly broadening its own field of talents. Witness the acrylics, the ABS polymers, the fluorocarbons, the LD and the HD polyethylenes, and even the polyacetals and polycarbonates scarcely with their umbilicals healed are sprouting unanticipated capabilities. This embarrassment of materials is not listed as an excuse, only as one of the facts industry is faced with. Another fact is that regardless of a company's size there is a limit as to how many hands can be assigned to bring on a new polymer, as to how much money can be allocated, and as to how long the development work can be carried out timewise to assure an economic return on a new, properly designed, and soundly merchandised polymer. These are hard questions to answer, but a rule of thumb could be written that for each major polymer to be brought to the market place an expenditure of 5–7 years in time and 15 million in dollars can be anticipated.

For specific comments on the limitation of plastics: long term performance data admittedly are needed; additional time is necessary to fit these newer materials into a wider variety of test schedules. The relatively high and sometimes variable thermal expansion can be markedly controlled according to preliminary tests through use of fibrous rather than pulverized fillers. Most of the newer polymers have no mechanical plasticization and, consequently, have no loss by migration, leaching, or volatilization. Similarly, effects of moisture pickup upon strength properties and dimensions are at a minimum in the newer, rigid polyolefins, for example.

Weatherability improvements have been made for these products, stepping up outdoor exposure endurance from a matter of a few months to several years. Still further progress can be expected. Batch-to-batch variation will always be with us whether it be in crops, metals, or plastics. Large scale blending, upwards of 50,000 lb per finished lot, help reduce significant fluctuation. Desired characteristics such as nonflammability, good scratch resistance, resistance to thermal, electrical, and nuclear exposure are inherent properties of materials. Certain applications in industries such as automotive, will put a premium upon specific properties and discount others. Economic justification, either in volume consumption or cash on the barrel head, can pretty much guarantee supply of a desired material. We can point to such laboratory exotics as 3-methyl, 1-butene (branched a-olefin) with a melting point of 310 C (590 F) and specific gravity of 0.91; or polyvinyl cyclohexane (alicyclic a-olefins) with a melting point of 700 F and a density of 0.95 and list their desirable properties for your consideration; but the ultimate market cost and demand would have to be spelled out in considerable detail before large-scale evaluation by a manufacturer could be undertaken. The point we are making is that the tools, and possibly the materials, are available to tailor-make polymers through stereospecific catalysis, graft- and block-polymerization, to name some recent techniques. But before the chemical industry goes down this path, the requirements of the automotive industry have to be made known, the volume and price structures have to be clarified. In other words, the two industries are suffering from a lack of communication in an area of vital importance to both.

Let's look at some specifics in the paper. Blow-molding developments can be placed in the following perspective: A year ago there was no domestically available blow-molding equipment, today there are eight companies out selling these tools. A year ago there probably were 12–15 companies blow-molding plastic, today there are at least 70 locations for such operations. Two blow-molded applications having some potential in automotive are: (1) a high-density polyethylene gasoline tank float: weight of 0.017 lb, raw material cost (only) 0.7¢; and (2) a brake fluid reservoir: weight 0.065 lb, raw material cost 2.5¢. These basic figures have undoubtedly caused a number of scratch pad calculations.

In the field of high density polyethylene, the design and use possibilities of very high molecular weight polymers in the region of several million average weight molecular weight (rather than the several hundred thousand normally produced) encourage some stimulating thinking. The polymer has excellent fatigue and shock absorption characteristics. Already these properties are being utilized in such items as textile picker blocks.

The dimensional stability and abrasion resistance of chlorinated polyether is being studied for automotive usage. One basis for this interest is a case history of a wobble plate in a water meter. Here the chlorinated polyether wobble plate showed no measureable wear after a standard test exposure of pumping 100,000 gal of water at 220 F whereas other plastics, as well as brass, showed noticeable wear. Incidentally, this item in production is molded to a $-0+0.001$ in. on the 3 in. diameter disc.

Finally, let's dream a bit on the basis of the subject matter of this paper. We have inherent advantages of freedom from corrosion, noiselessness, and improvements in reducing unsprung weight due to lightness of these plastics. They are reasonably rigid, have good tensile strength and substantial heat resistance. Some of these properties can be further improved by orienting these polymers as they are fabricated, for example, by stamping, whereby the plastic is drawn at a temperature below its melting point. Strength approaching mild steel can be attained in such structures. The utilization of thermoformed sheet structure and pumping in foam-in-place plastics for roofing, door panels, trunk lids, floors, and the like, should be possibilities.

While the possibilities are logical, not all will prove to be economical. However, such papers as this one and such meetings as we have here today can't help but spark some combined thinking to the advantage of both the automotive and the plastic industries.

ELECTRIC PROPULSION

—measurements with a small thrust plasma generator

B. Gorowitz and B. W. Harned†

Aerosciences Laboratory
General Electric Co.

IN PRELIMINARY studies concerning the use of plasma acceleration for propulsion purposes, a series of measurements has been made of velocity and momentum of plasma bursts both in single-shot and pulsed operation over a range of pressures and input energies for various magnitudes and directions of magnetic field at the gaseous discharge. The effect of magnetic field on acceleration of the plasma and its contribution to the momentum transferred to a ballistic pendulum are discussed. Extrapolation has been made of power input for the small thrust obtained to that required for a one pound thrust over the range of pressures and efficiencies investigated.

This paper is based on work performed under the auspices of the Ballistic Missiles Division, U. S. Air Force, Contract No. AF 04(647)-269.*

SEVERAL TYPES of plasma accelerating devices have been proposed[1] for possible incorporation in a small thrust electric drive for space propulsion or orientation control. Practically all involve the use of magnetic coupling to the ionized, but electrically neutral, material comprising the plasma to increase exhaust energy by electromagnetic repulsion. Some devices use electrodes to produce the plasma with the magnetic field normal to the discharge so that a Lorentz force exists in the direction of plasma exhaust. The T-tube to be described herein is one of this class. Another type of coupling is one with the magnetic field directed along the exhaust axis, induced currents being set up in the plasma by a changing magnetic field and causing repulsion in accordance with the Lenz law. This latter method has been used without electrodes[2] with rapid, high current discharge through a coil producing the plasma and the necessary field conditions. Thus, if one considers a magnetic field B, plasma repulsion may be obtained with both a high B and a high (dB/dt). Materials for use as fuels in such systems are those which ionize easily, since the degree of coupling of the magnetic field depends on the conductivity of the plasma.

Most of these plasma devices require large currents for production of both high conductivity in the plasma and large magnetic fields. The present means of obtaining these currents is to release capacitor-stored energy to the electrical circuit within a very short time, allowing thousands of amperes to flow during peak current surges. These circuits may or may not "ring" at a characteristic frequency depending on the amount of resistance of the associated circuit. Obviously, the lower the resistance, the greater is the likelihood of ringing, but at the same time the higher the peak currents and magnetic fields. In such circuits the lower the associated inductance, the higher is the ringing frequency and, consequently, the higher are the peak currents in the first portion of the discharge when the greatest action is expected. Very low inductance is desirable in both means of magnetic acceleration, but especially so with the (dB/dt) type. Propulsion is visualized with such devices because of the extremely high velocities (and hence low masses for any given rate of momentum transfer), which can be realized with the aid of the magnetic coupling. Basically, the method is to introduce the fuel to a chamber, convert it to a plasma, and apply the magnetic field. The process can be a continuous one, but the attainment of high magnetic field strengths presently is difficult with other than capacitor discharge circuits. Consequently, plasma accelerators with such energy storage circuits must necessarily be pulsed. It must be appreciated that the power requirements for propulsion are very large for even small thrusts; that is, for thrusts of the order of 1 lb, the power supply must be capable of delivering about a megawatt of high voltages. Re-

† Now at Flight Propulsion Division, General Electric Co.

* Paper presented at SAE National Aeronautic Meeting, New York, April 3, 1959.

[1] For example, the button and rail sources discussed in "Plasma Motors," by W. H. Bostick, *Journal of Astronautical Sciences*, Vol. 5, February, 1958, p. 39. Also, the toroidal source discussed in "Description of Propulsion Device which Employs Magnetic Field as Driving Force," by R. M. Patrick, AVCO Research Laboratory Report 28, May, 1958.

[2] "Experiments Using Hydromagnetic Shock Tube," by V. H. Blackman. Paper presented at Second Magnetohydrodynamics Symposium, Lockheed Research Laboratory, Palo Alto, Calif., December, 1957.

quirements for circuit components are, therefore, great, especially when one considers the necessity of these components handling such large powers over long periods of time because of the rather low thrusts.

We should like to present data on operation of a very low efficiency device as representative values for preliminary work in investigation of the feasibility of this type of electric propulsion. We shall attempt to make clear the areas of inefficiency and to point out possible means of improvement. We have made measurements of velocity of bursts of air plasma and momentum available from these bursts for both single-shot and pulsed operation. The ideal conditions of very low pressure exhaust have not yet been attempted; alternatively, we have used the easier experimental approach of maintaining exhaust tank and discharge chamber at the same pressure before production and acceleration of the plasma. The errors involved in this technique will be discussed later. We have measured for this system the efficiency of conversion of electrical to mechanical energy over a range of energies and ambient pressures for different magnitudes and directions of applied magnetic fields at the discharge.

Circuit Details

Our T-tube is essentially that described by Kolb[3,4] and Kash[5] modified somewhat to provide varying magnetic fields. It consists of a 39-mm ID pyrex tube in the form of a T, two arms of which house brass electrodes separated by a 30-mm gap, while the third arm is used for exhaust of the plasma set up in discharge. Fowler[6] showed that capacitor discharge through such a tube causes rapid expansion of the heated gas which sets up shock waves in the exhaust portion of the tube. Kolb[2] placed the return lead of the electrical circuit behind and parallel to the T-tube discharge so that the high currents set up large magnetic fields at the discharge to aid in the expansion process. This connection permits a force on the ionized material in the gaseous discharge as long as current is flowing in the circuit, but suffers from the disadvantage that, as displacement of the discharge is increased, the repulsive force is decreased because of the s^{-1} dependence of the magnetic field from the location of the return lead. Alternative return circuits have been made by substituting 1,- 2-, and 4-turn coils arranged to produce suitably directed magnetic fields at the ionized material. A photograph of the assembly used is shown in Fig. 1. Our circuit in simple form is as shown in Fig. 2, where the battery of potential difference V represents a d-c power supply, R a charging resistance, C the capacitor, S a triggered spark gap switch, and L the load which is the low pressure spark gap of the T-tube. This circuit for dissipation considerations can be drawn more simply as Fig. 3, where R_L represents the capacitor discharge circuit and R the charging resistor. It is immediately seen that high current drain from the power supply requires an undesirable loss of power in this charging resistor, whose only function is to maintain the peak and average current delivered by the power supply within specifications.

The capacitor C is made up of two GE model 14F630 5-microfarad 20-kv capacitors in parallel, and the resistor R is made up of 12 200-watt units in series-parallel arrangement to provide to total of 1200 ohms. Two d-c power supplies were used for capacitor charging: a 10 kw, 5 kv-2 amp supply served for pulsed operation, while a 20 kv-50 ma supply allowed our single-shot measurements to be extended to 10 kv; both units were made by Precise Measurements Co. Pulsing frequency was usually 20/sec at 4.5 kv, with an average current drain of 0.89 amp delivering 950 watts to the charging resistor and 2000 watts to the capacitor. The ringing frequency of the discharge circuit varied from 37–72 kc/sec with discharge circuit resistance and inductance values varying from 0.05 to 0.12 ohms and 0.57 to 1.84 microhenries, respectively. Current surges on discharge at 10 kv operation were as high as 34 kiloamperes with magnetic field strengths at the discharge of about 4000 gauss.

The spark gap switch S shown in Fig. 2 is made up essentially of two copper pieces rounded to about 1-in. radius. Along the axis of the high voltage electrode is located a tungsten trigger electrode, insulated by a ceramic tube, the end of which projects slightly beyond the tip of the copper. A high voltage pulse is applied to this trigger electrode from the circuit of Fig. 4, causing the switch gap to break down and the capacitor to discharge. This switch is necessary for the high voltages and pressure range over which we operated because the T-tube gap breaks down at too low a voltage otherwise. For pulsed operation we have had much difficulty in coaxing the switch to follow the trigger pulsing. Some improvement was gained in response of the switch when the electrodes were hollowed out and grooved radially at the tips. The trigger pulse is obtained from a pulse transformer in the plate circuit of a 2D21 thyratron which discharges a capacitor through the transformer primary when fired by a pulse applied to its grid circuit. Both a Hickok sine-square wave generator and a Tektronix pulsing circuit have been used to provide this initial triggering, the latter being somewhat more useful because of the very low repetition rates obtainable. The switch and trigger circuit are modifications of those of the AFCRC Lovotron.[7]

The exhaust chamber into which the accelerated plasmas are driven consists of a 12-in. OD brass cylinder, 12 in. high, on top of which is placed a 20 in. high bell jar. Three 5-in. glass windows in the cylinder allow added observational ability. This chamber is pumped out continuously by a rotary-type vacuum pump. Sufficient volume is enclosed so that the chamber tends to act as a reservoir decreasing pressure fluctuations even during pulsed operation of the T-tube. The fuel, air, is admitted continuously at the low voltage electrode through a Veeco variable leak. Pressure in the chamber is monitored by a Stokes-McLeod mercury gage.

The first variation of return lead to enhance the

---

[3] "Production of High-Energy Plasmas by Magnetically Driven Shock Waves," by A. C. Kolb. *Physical Review*, Vol. 107, July 15, 1957, pp. 345–350.

[4] "Magnetically Driven Shock Waves, Experiments at U. S. Naval Research Laboratory," by A. C. Kolb. From "Magnetohydrodynamics," pp. 76–91. Pub. by Stanford University Press, 1957.

[5] "Magnetically Driven Shock Waves, Experiments at Lockheed Missile Systems Division," by S. W. Kash. From "Magnetohydrodynamics," pp. 92–98. Pub. by Stanford University Press, 1957.

[6] "Final Report, Research on Radiation Transients in Gas Discharges," by R. G. Fowler, University of Oklahoma, December, 1952.

[7] "Lovotron — Low Voltage Triggered Gap Switch," by Cullington, Chace, and Morgan. AFCRC TR-55-227, Air Force Cambridge Research Center, September, 1955.

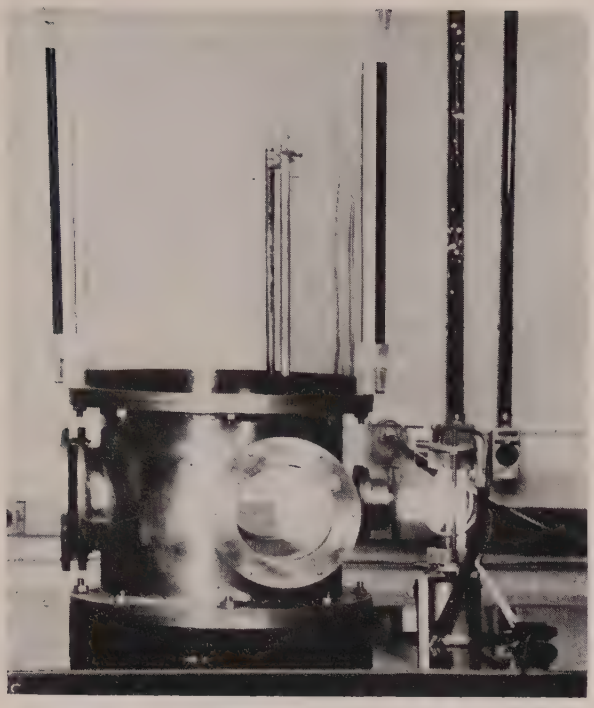


Fig. 1 — T-tube accelerator-vacuum chamber assembly

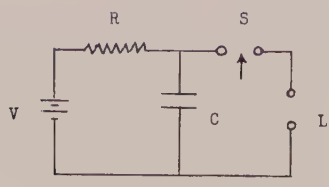


Fig. 2 — Discharge circuit

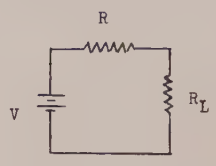


Fig. 3

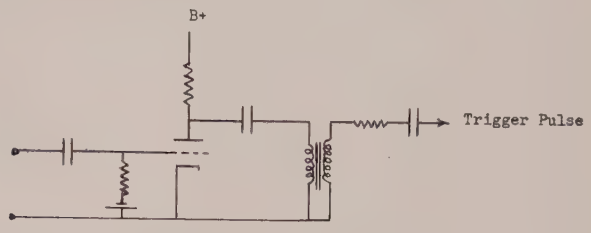


Fig. 4 — Pulser circuit

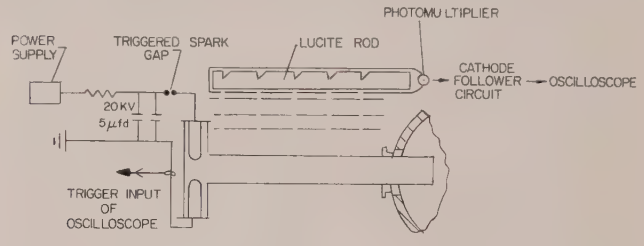


Fig. 5 — Schematic velocity measurement apparatus

magnetic field at the spark discharge was that of a two-turn coil, one turn of about 7.5 cm diameter on either side of the ionized stream, so that a more uniform field could be obtained over that of the single strap in back of the discharge. During pulsed operation, at low voltages, a marked displacement of the radiating gases could be seen as a result of this coil, and more intense radiation was visible in the exhaust chamber. The second variation was that of a 4-turn coil, two turns of 7.8 cm diameter on either side. Other variations were: (1) the use of a longer lead of ¼-in. copper tubing away from the discharge to make negligible the repelling magnetic field; (2) the use of ¼-in. copper tubing for a return lead at the rear of the discharge but located at the same distance as the 1-in. brass strap initially used; (3) the use of a single-turn coil at the rear of the discharge with its magnetic field parallel with the direction of the exhaust; (4) the single turn coil as in (3) but rotated 90 deg so that a plane through its enclosed area intersected the electrodes; (4) the use of a strap with greater separation from the discharge; and (5) a 2-turn coil with one loop as in (3) and the other loop in front of the discharge so that a more intense and uniform field was created along the direction of the exhaust. It was not possible to match circuit inductance or resistance with these various return leads so that a range of frequencies and currents resulted for a given capacitor voltage. The use of coil configurations in the return lead is a deliberate introduction of inductance in the discharge circuit which admittedly sacrifices high initial peak currents for more intense and uniform magnetic fields. When this was originally done with the pulsed circuit, no means were available to measure plasma velocity, and it was believed that the coil was much more efficient under all operating conditions. The data to be presented here show this not to be the case, particularly at high input energies. Rotation of the magnetic field direction through 90 deg so that it is directed along the exhaust causes a deflection of the plasma at right angles to the exhaust and hence reduces considerably the deduced conversion efficiency of electrical to mechanical energy.

Method of Measurements

Velocity—We have used the photomultiplier technique of velocity measurement, with the assumption that transit of luminous material past an observation port represents adequately the movement of the accelerated material. Kash[5] has shown for the T-tube that a discrepancy exists between photomultiplier and microwave measurements, depending on photomultiplier sensitivity, and that error is inherent in this type of measurement depending on degree of luminosity. We have used only a single detector, connected at the end of a lucite "light pipe," into which light from five equally spaced apertures (4 cm) passes in a normal direction from the axis of the exhaust tube (Fig. 5). The aperture system is sufficiently stopped down to confine the light cone angle to less than 5 deg. The photomultiplier was connected to a Tektronix dual-beam oscilloscope through a cathode follower circuit whose rise time is 0.1 microsec or better. The oscilloscope was triggered by a voltage pulse induced in a loop around the return lead to the capacitor, and this waveform was also displayed on the oscilloscope in order to provide information about timing with respect to the discharge current maxima and minima. Due to successive plasma bursts exhausted past the light slits, it was necessary to close all slits

but one when making measurements, because overlap of pulses caused confusion as to which pulse belonged to which aperture. Good reproducibility of start of the oscilloscope trace assured the usefulness of the method, which, for any one set of conditions, required five different discharges to allow timing of arrival of the accelerated gases at each of the five apertures to be determined. The method is useful in its ability to time successive pulses and thereby compare velocities of successive accelerated gas volumes. Velocities were computed from slopes of the distance-time curves drawn for the arrivals of the first and second pulses at the five stations, an example of which is shown in Fig. 6. A typical oscilloscope trace is shown in Fig. 7.

Momentum—The ballistic pendulum method has been used to determine the amount of momentum available from the accelerated gases. A disc was suspended by two strings of ordinary thread within the exhaust chamber at the output of the T-tube exhaust arm. Two discs were used for these measurements: the one was a circular plastic lid of an icebox dish, 3.75 in. diameter and 11.5 grams mass; the other was made by mounting aluminum shim stock on a plastic frame to form a 3.5-in. disc of 3.2 grams mass. Period of the pendulum was calculated to be about 1.3 sec and measured to be the same value. For our purposes, the system has been treated as a simple pendulum, so that the momentum of the pendulum (MV) is used to calculate the momentum of the plasma (mv) which moves it. V of the pendulum is calculated from the maximum angle of swing for single-shot measurements, using the usual energy relationship to provide $V = \sqrt{2gL\,(1 - \cos\theta)}$ where L is the length of the string. Measurements of angle were obtained from a scale beneath the pendulum calibrated directly in angular displacement, and were obtained by visual observation only, the average of three displacements used for the calculations.

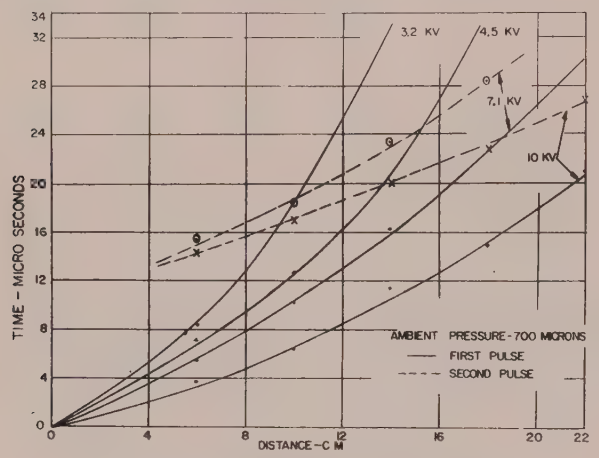


Fig. 6 — Photomultiplier time-distance data

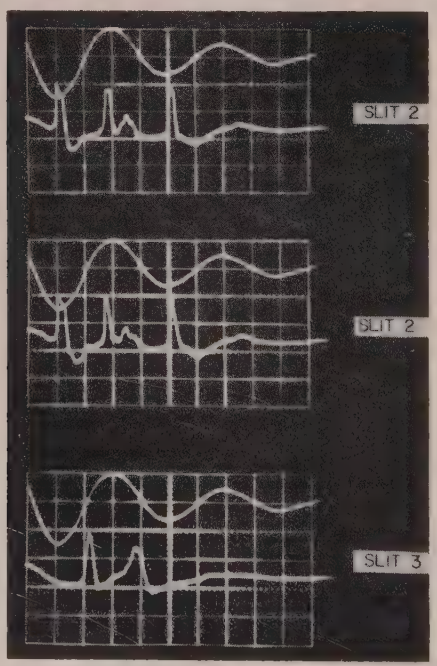


Fig. 7 — Oscillogram of photomultiplier and pickup coil signals (7100 volts, 700 microns pressure, 2-turn coil with field perpendicular to discharge)

Identical diameter but different length T-tubes were used in these series of measurements. A 12-in. exhaust length was needed for the velocity series, while a 6-in. exhaust tube was used for the momentum series. In this fashion we felt that more reliable values of plasma velocities could be obtained for the distance at which the pendulum was located. The velocity of the plasma decreases with distance s from the electrodes very nearly as $s^{-0.5}$, as predicted by Harris[8] and experimentally verified by Kolb[3] and the present authors in unpublished work with another type of plasma accelerator.[9]

For pulsed operation one may use the angular scale reading as a direct measurement of thrust, provided that the period of the pendulum is much longer than the time between pulses. The thrust then is given by $F = Mg \tan\theta$, where M and g are mass of the pendulum and gravitational acceleration, respectively. An aperiodic bobbing of the pendulum has been observed during pulsed operation, making such measurements difficult because of rather wide swinging. The cause of this bobbing has not been investigated thoroughly, but a number of factors may contribute to the action. Primarily, it is believed that the thrust varies because of decreased mass per shot as a result of heating at the electrodes. It is also believed that outgassing of the electrodes occurs at the start of the pulsing operation and decreases as the electrodes become heated. A continuous decrease of the outgassing would be expected to cause a gradual decrease in height of the pendulum with steady operation of the capacitor switch, and such action was noted. In fact, momentum and thrust measurements given later as extrapolated values from single shot discharges are somewhat higher than those measured for pulsed operation, the discrepancy increasing with time of operation by as much as 50%. An electrostatic factor also may cause a bobbing effect as a result of operation of the T-tube, chamber, and pendulum, which are mutually electrically insulated.

We should like to point out that a considerable error may result in thrust measurement by the static method of angular displacement of the pendulum, because of the increased distance from the end of the tube of impact of the plasma. The error thus encountered involves a decreased velocity from that

[8] "Exact and Approximate Treatments of One-Dimensional Blast Wave," by E. G. Harris. Naval Research Laboratory Report NRL-4858, 1956.

[9] "Production of High-Velocity Shocks," by Vernal Josephson. *Journal of Applied Physics*, Vol. 29, January, 1958, pp. 30–32.

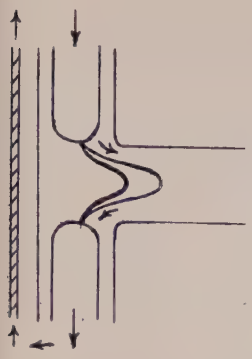


Fig. 8 — Displacement of gaseous discharge due to magnetic field of current in return strap

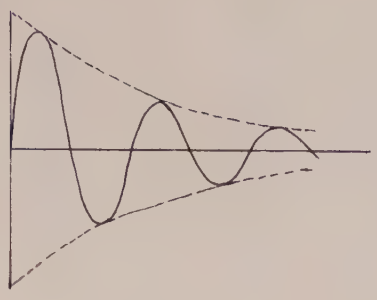


Fig. 9 — Induced voltage curve

calculated, and a decreased mass results from diverging of the plasma beam. We have not taken such errors into account, believing that the overall error is in our favor. The fact that single-shot momentum measurements, extrapolated to provide thrust values for pulsed operation, are higher than the measured pulsed values is good evidence for this expected error.

Energy of Plasma and Efficiency of System—The product of pendulum momentum for single-shot discharges and plasma velocity at the tube exhaust distance must equal twice the plasma energy if conservation of momentum is assumed. Efficiency of conversion of capacitor stored energy to mechanical energy is thus a ratio of plasma energy $1/2\ (MV)v$ to $1/2\ CE^2$. In order to calculate power dissipated in the charging resistor during pulsed operation, one must find the effective heating current delivered during the charging time. This is done by the usual integration process and the power is given by:

$$I^2_{eff}\,R = \frac{1}{T}\left(1/2CE^2\right)\left[1 - e\frac{-2^r}{RC}\right] \qquad (1)$$

where: T = Period of pulsing

Power input from the power supply is thus a sum of $I^2_{eff}\ R$ and $1/2\ CE^2f$ where f is the pulsing frequency of R is the charging resistor, and amounts for all practical purposes to twice the capacitor power.

Inductance and Resistance of Discharge Circuit— The induced voltage in the loop around the return lead of the discharge circuit allows display of a damped waveform, the decrement of which is given by $e^{(r/2L)t}$, where r is the resistance of the capacitor discharge circuit and L the associated inductance. The period of the ringing can be determined directly from the waveform. When the gaseous discharge is shorted out, the same type of induced voltage waveform produces inductance and resistance values exclusive of the T-tube, which in turn allow representative values to be obtained of the T-tube portion of the circuit.

The resistance of the return leads can be calculated approximately by using the physical dimensions and the skin depth relationship.[10] Since the electrode resistance of the T-tube should be relatively constant, one can arrive at representative values of the actual gaseous discharge resistance under various conditions. This value then allows an idea of the amount of power going into the gaseous portion of the electrical circuit.

Maximum Currents and Magnetic Fields — The current in a ringing circuit is given approximately by

$$I = V_o\sqrt{\frac{L}{C}}e^{-R/2Lt}\sin\sqrt{\frac{1}{LC}}\,t$$

with maximum values when $\sin\dfrac{1}{LC}t = 1$; the first maximum then is given by:

$$i_m = V_o\sqrt{\frac{C}{L}}e - \frac{\pi R}{4}\sqrt{\frac{C}{L}} \qquad (2)$$

where R and L represent the resistance and inductance of the discharge portion of the circuit. These values of current for the various values of voltage V to which the capacitor is charged then determine the values of magnetic field which have been calculated in accordance with the geometry of the return leads employed.[11] The process of repeated installation of the various return lead configurations probably resulted somewhat in altered coil dimensions, so that it is to be realized that these field values are only approximate. In addition, no account was made of the field due to the gaseous portion of the circuit, that is, the mutual inductance, the idea being merely to determine an estimate of the magnetic field intensity at the plasma due to the calculated currents in the return leads.

Results

As might be expected, successive current peaks produce displacements of the ionized material making up the discharge, and these are visible for a single observational port with a photomultiplier assembly. For stations close to the electrodes one might imagine a ribbon of material kinking out as in Fig. 8 past the observation points. The distance of displacement depends on input energy, and, presumably, with sufficient energy one might obtain displacement the total length of the T-tube exhaust section. This kinking is the actual path of the current, so that with a ringing circuit with changing directions of current as shown in Fig. 9, one might expect that any displacement outward must be followed by a displacement back to the electrodes. Hence, with fine enough resolution, each station may see two pulses for each current peak, provided, of course, that the luminous material is confined well in size. Actual shape of the kinking after it starts will depend on conductivity of the gas, and it should be noted that magnetic fields due to currents on opposite sides of the ribbon will tend to maintain the kink once it starts. Some indications of a returning current were obtained on our traces, but one cannot say they are conclusive of the above. It is possible that a parallel path is set up for current conduction nearer the electrodes once the discharge has been displaced, so that a break in the kink is made and another kink is started simultaneously at much lower energy. Even though the discharge may not reach all stations, one may expect that it can act as a piston which sets up a

[10] "Transmission Lines, Antennas and Wave Guides," by King, Mimno, and Wing. Pub by McGraw Hill Book Co., New York, 1945.

[11] "Electric and Magnetic Fields," by Attwood, Chap. 10. Pub. by Wiley and Sons, New York, 1949.

Table 1 — Operating Conditions

Return Lead		Strap, 1 in. Separation				Single Turn, either side			
Energy Input, joules	P	500	250	100	50	500	250	100	50
1. Velociety measurements, 1st pulse. cm/microsec 18 cm from electrodes	200	1.75	1.15	0.41	—	1.65	0.99	0.67	0.29
	400	1.17	0.61	0.32	—	1.22	0.73	0.36	0.13
	700	0.75	0.52	—	—	0.79	0.56	0.23	—
	1,000	0.68	0.40	—	—	0.65	—	—	—
2. Momentum measurements, dyne-sec	200	345	145	61	38	351	144	69	23
	400	409	173	85	61	408	170	96	80
	700	454	211	101	82	486	198	145	96
	1,000	524	262	125	96	621	353	171	125
3. Plasma energies, joules	200	30.2	8.3	1.25	—	29	7.7	2.3	0.6
	400	24	5.3	1.35	—	25	6.2	1.75	—
	700	17	5.5	—	—	19	5.55	1.65	—
	1,000	18	5.2	—	—	20	—	—	—
4. Efficiencies % of capacitor input energy	200	6.1	3.2	1.25	—	5.8	2.8	2.3	1.2
	400	4.8	2.1	1.35	—	5.0	2.5	1.8	—
	700	3.4	2.2	—	—	3.8	2.2	1.7	—
	1,000	3.6	2.1	—	—	4.0	—	—	—
5. Mass delivered, micrograms/pulse	200	197	126	—	—	212	145	101	80
	400	350	283	—	—	335	233	—	—
	700	605	405	—	—	612	353	—	—
	1,000	772	653	—	—	955	—	—	—
6. Thrust, calc. using momenta above at freq. of 20/sec ounces	200	0.25	0.10	0.04	0.03	0.25	0.10	0.05	0.02
	400	0.29	0.12	0.06	0.04	0.29	0.12	0.07	0.06
	700	0.33	0.15	0.07	0.06	0.35	0.14	0.10	0.07
	1,000	0.38	0.19	0.09	0.07	0.45	0.25	0.12	0.09
7. Input power per lb thrust, above efficiencies f = 20/sec megawatts	200	0.64	0.80	0.80	0.53	0.64	0.80	0.64	0.80
	400	0.55	0.67	0.53	0.40	0.55	0.67	0.46	0.27
	700	0.49	0.53	0.46	0.27	0.46	0.57	0.32	0.23
	1,000	0.42	0.42	0.36	0.23	0.36	0.32	0.27	0.18
8. Fuel exhausted in lb/hr per lb/thrust	200	2.02	3.15	—	—	2.24	3.59	—	—
	400	3.83	5.85	—	—	2.93	4.94	—	—
	700	4.72	6.70	—	—	4.50	6.28	—	—
	1,000	5.18	8.75	—	—	5.40	—	—	—

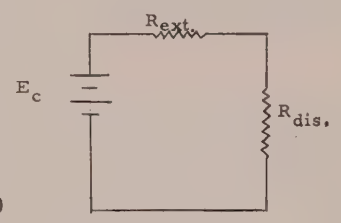


Fig. 10

shock wave in gas further down the tube. For our case, the times of arrivals of many of the pulses were much too large for the actual discharge to reach the furthermost stations during the current flow, it not being expected that any such breakdown path could be maintained during the time of current reversal because of ion recombinations.

On current reversal one expects a second kink less energetic than the first because of the reduced current value at the second maximum, this followed by as many kinks as rate of energy dissipation allows. The capacitor energy is dissipated logarithmically over several cycles, but so far as magmetic coupling is concerned, one may safely say that the greatest effect is during the first two current peaks or one cycle of ringing. That the second half cycle is effective is seen in plots of distance-time for the second pulses. The velocities of the second pulses are higher than the first, so that for the more einergetic discharges the second advancing shock can catch up with the first in the gas further down the tube. Velocities have not been calculated for third and successive pulses because our sensitivity was not great enough to record these.

The existence of the second highly energetic pulses introduces a great deal of error in our calculations of plasma energy, accelerated material mass, and efficiency of conversion of energy. Some of the energy dissipated in the first half cycle goes into vaporization of the electrodes, probably through the sputtering process, and some of this vaporized material makes up that which hits the pendulum in the exhaust chamber. It takes a few microseconds to get this material into the stream,[12] so it is to be expected (although not measured by us) that the second half cycle is effective in moving this metallic vapor. We have no way presently of time resolving the accelerated masses with respect to electrode current, but the measurement in principle ought to be done fairly easily by means of monochromatic separation and photomultiplier detection of radiation emitted at the exhaust of the T-tube. We have not attempted to weigh the material deposited on our pendulum, partly because a great deal of reflection occurs, as seen from deposited material on the inside of our exhaust chamber. To be sure, such a value would provide a minimum value for energy calculations, but we believe that such is also true by using velocity of the first pulse which averaged about 20% lower than that of the second.

In Tables 1 and 2 are presented our data for the various conditions of operation. We should like to point out that we have stated three significant figures, more for consistency in calculation than actual measurement accuracy, and hasten to point out that our concern has been more for order of magnitude in comparison than the actual numbers. From the data then, we should like to indicate a

[12] "Short Period Investigations in Spark Advance," by D. W. Steinhaus, H. M. Crosswhite, and G. H. Dieke. *Journal of Optical Society of America*, vol. 3, April, 1953, pp. 257–270.

Table 2 — Operating Conditions

Return Lead	F	L	R_{ckt}	R_{dis}	$\frac{R_{dis}}{R_{ckt}}$	L_{dis}	I^1_{max}	B^1_{max}	V^1	V^2	MV	Eff
1. No field	52.6	0.913	0.054	0.024	0.45	0.417	28.8	—	1.51	1.51	244	3.7
2. Strap, 2 in. sep.	64.5	0.608	0.057	—	—	0.112	33.9	—	2.0	1.63	—	—
3. Strap, 1 in. sep.	66.7	0.569	0.063	—	—	0.073	33.9	—	1.75	1.75	345	6.1
4. Single turn, either side	51.2	0.961	0.095	0.064	0.68	0.465	25.3	2,750	1.65	1.64	351	5.8
5. Single turn, front & back	51.2	0.961	0.074	0.042	0.57	0.465	26.7	3,850	0.93	1.05	109	1.0
6. Double turn, either side	37	1.845	0.117	0.083	0.71	1.349	18.8	4,300	1.26	1.61	345	4.4
7. Single turn, back, field parallel	54	0.866	0.061	0.031	0.51	0.370	28.9	—	1.16	1.5	—	—
8. Single turn back, field perpendicular	55.5	0.820	0.076	0.046	0.61	0.324	28.3	—	1.44	1.9	—	—
9. 0.25-in. tube as strap	64.5	0.608	0.060	0.032	0.53	0.112	33.5	2,750	1.80	2.05	486	7.4
10. Discharge shorted	71.4	0.496	0.026	—	—	—	—	—	—	—	—	—

All data for 10 kv on capacitor (500 joules input) and 200 microns air pressure, single-shot work

Symbols: F = Frequency in kc/sec
L = Discharge circuit inductance in microhenries
R_{ckt} = Discharge resistance in ohms
R_{dis} = T-tube resistance = $R_{ckt} - 0.026 -$ (calculated resistance of lead)
L_{dis} = T-tube inductance in microhenries = $L - 0.496$
I^1_{max} = Calculated first current maximum in kiloamp
B^1_{max} = Calculated first maximum field intensity at discharge using I^1_{max} and geometry, gauss
V^1 = Measured velocity of first pulse 18 cm from electrodes, cm/ /u sec
V^2 = Measured velocity of second pulse 18 cm from electrodes, cm/ /u sec
MV = Measured momentum of pendulum due to 10 kv discharge in dyne-sec
Eff = 100 × ratio of plasma energy to $\frac{1}{2}CV^2$, using v^1 for calculation, %

few qualitative facts.

1. The originally supposed increased efficiency of the coil configuration in the return lead over the back strap efficiency is true only for the lowest energies significantly so only at the lowest pressures. At high energies, the strap appears more efficient, the added energy in the magnetic field apparently not as efficiently returnable to the plasma. That our acceleration is proportional to the value of B, rather than the value of $\frac{dB}{dt}$ is seen in comparison of the two cases of the use of the 2-turn coils, one field parallel and the other perpendicular to the discharge.

2. The increased magnetic field at the discharge tends also to increase the calculated resistance of the discharge, as deduced from the induced voltage waveform, and shown by $R_{dis.}$ If the capacitor is considered as a battery E_c as in (Fig. 10), with the circuit resistance given by R_{ext} and that of the gaseous conductor with mutual coupling to the magnetic field given by $R_{dis.}$ then for capacitor energy dissipation purposes, one can say that the ratio of $R_{dis}/(R_{dis} + R_{ext})$ of this energy actually gets to the gaseous conductor part of the circuit. The highest values of this ratio occur with good magnetic field coupling to provide the Lorentz force in the desired direction. If an "efficiency of discharge" is calculated using the designated efficiencies of capacitor input energy and the above ratio (thus effectively calculating (energy out)/(energy in) for the gaseous conductor), one finds by far the best efficiency (about 14%) at high energies at 200 microns pressure in the ¼-in. tubing used as a return lead, while a wider strap or single turn coil on either side proves to be only 60% as good (about 8%).

3. The fact that magnetic field coupling tends to increase resistance indicates that resistance actually varies with current in the circuit. To test this, we have calculated resistance for the four voltages for the single-turn-either-side case and find that there is a tendency for the resistance to reduce as input energy is reduced. An average of several decremental curves was taken, and a decrease of about 20% is indicated in circuit resistance as energy decreases from 500 to 50 joules. This appears as a 30% decrease in gaseous conductor resistance. These changes were not considered in calculation of the values given in the tables.

4. A remarkable increase in mass delivered is seen for the higher pressures at the two input energies listed. One can make an estimate of the amount of gas in the tube which is available for delivery to the pendulum (that is, the volume of the exhaust tube), but the increased pressure cannot account for these large mass increases. The extra amount must be due to increased vaporization of the electrodes, which if it is a sputtering process, ought to depend on the number of ions striking the cathode surface. The ratios of the masses are somewhat less than the corresponding ratios of gas densities, but one would expect increased losses at higher pressure because of increased diffusion to the walls during exhaust.

5. Since the velocities are not directly comparable because of varying frequencies in the circuit for the different return leads, one might compare ratios of current maxima to the corresponding velocities measured to obtain an idea of how many kiloamperes are needed to attain a velocity of 1 cm/microsec. This set of ratios has been made for both the first and second pulses and, to no great surprise, one finds that the use of high magnetic field coupling provides the lowest current values while the highest values are for the no-field case and the cases where the magnetic field was directed along the exhaust direction.

6. When exhausting the accelerated plasma over the pressure range we have used, one must bear in mind that velocity values are quite dependent on distance from source, ambient pressure, and input energy. Consequently, efficiency values are also highly dependent on these quantities, and those derived herein should be considered as minimum values. The Harris relationship[8] predicts for such

velocities a dependence on energy, density, and distance as $W^{0.5}$, $\rho^{-0.5}$, and $s^{0.5}$. We have investigated this relationship thoroughly for another type of discharge and found some deviations from the 0.5 value of exponent. For this set of data we have a smaller range of parameters, but for the single-turn-either-side case, at 500 joules input, the pressure dependence is $p^{-0.58}$, while the distance dependence over the length of 18 cm at 700-microns pressure is $s^{-0.45}$. For 400-microns pressure a linear relationship is obtained for velocity versus voltage, substantiating the $W^{0.5}$ prediction over the relatively small range of input energies dealt with here. It is certainly to be appreciated that larger velocities would be measured if the exhaust tube were shorter, and presumably larger momenta would be observed because of smaller loss of vaporized material on the walls of the exhaust tube.

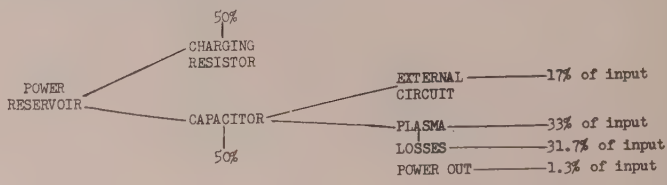
7. The calculations for input power per pound of thrust and rate of fuel exhastion per pound of thrust are based on the repetition rate of 20/sec and the momentum and efficiency values measured. A striking fact is revealed in the decrease in input power required at the higher pressures to produce the output one pound of thrust, especially when the efficiency of conversion goes down with the higher pressures. That this is true may be seen from the method of calculation which shows the input power relationship to be a conversion factor times the ratio of velocity to efficiency of energy conversion. The input power for the higher pressures is, therefore, seen to be reduced because the velocity efficiency ratio is reduced, undoubtedly due to the ability to deliver more mass through sputtering of the electrodes at the higher pressures. This simple relationship certainly points out that one cannot strive continually for higher velocities of the exhaust material without equal consideration for the efficiency of energy of conversion if minimum input powers are important. The use of auxiliary circuitry to provide sustained magnetic coupling to the plasmas for higher exhaust velocities must certainly be a high efficiency process to cause a reduction of this ratio.

Conclusions

Although this set of measurements is purely academic so far as simulation of flight conditions is concerned, it does indicate several areas of investigation for further work in electric propelling devices. The pressures we have used are high compared to the 10^{-12} mm mercury pressure or lower for space operation, and the losses due to exhaust of the plasmas into these high densities are apparent. One may term the collision losses of energy through ionization and diffusion a type of friction which would be very greatly reduced with vacuum exhaust, such that the velocity dependence on pressure and distance from the electrodes may be greatly changed. The problem remains as to which density to use for creation of a plasma for the particular gun mechanism selected. The T-tube experiment herein described indicates that higher densities are useful for an electrode-type gun because of the added sputtering as an efficient process of energy conversion. The containment of fuel at such a pressure prior to ionization may be difficult for an actual system.

From an electrical dissipation point of view, it would appear that one would try to make the effective resistance of the gaseous conductor as large as possible and minimize external circuit resistance in order that most of the capacitor energy appears in the plasma. The addition of a magnetic field coupling is a means of increasing this factor, but our data indicate that the energy conversion efficiency is not linear with the resistance increase. That the ¼-in. tubing type of return lead with its lower coupling should be more efficient than the 2- or 4-turn coil indicates that the processes of conduction and repulsion are complicated. The ratios of product of current and field intensity should give an idea of the corresponding ratios of velocity, but one does not see in our data any such correlation.

A partition of power for one of our circuits is shown below as typical for pulsed operation at 4.5-kv power supply voltage.



The 50% loss in the charging resistor and the 17% loss in the discharge circuit external resistance are definitely points for engineering attention. With high current capacity power supplies and low resistance connections of the circuit and capacitor switch these losses should be reduced considerably in future operations. Delivery of more than 1.5% of overall energy input to the exhaust undoubtedly will improve with vacuum exhaust, but the amount of improvement is difficult to foresee at the present time from the data of this experiment. A great deal of energy is lost in radiation from the hot gases during transit, and a choice of fuel may be made which will reduce this loss. A d-c axial magnetic field may be useful to reduce wall losses, as suggested by the experiment of the New York University group.[13] Much energy is certainly conducted to the electrodes and electrodeless production of plasma is the obvious means to avoid this loss, provided that sufficient density of plasma can be created.[14] Operation at higher power supply voltages does increase efficiency, as our data show, so that 10-kv operation should raise the output figure to over 3%.

The overall efficiency of this type of accelerator is disappointingly small at the present time, comparatively speaking, but it is felt that the process of magnetic acceleration of plasmas is somewhat better understood through this work, sufficient to encourage further research and development.

Acknowledgments

We are grateful for many helpful suggestions from a number of our co-workers, particularly those of Dr. J. Farber, whose continued interest in our experimental procedures has led in part to the completeness of this report.

[13] "Experiments on Supersonic Plasma Flow Along Magnetic Fields," by Wetstone, Finkelstein, Ehrlich, and Livingston. Paper presented at meeting of American Physical Society, New York, January, 1959.

[14] For example, the abstract of theoretical analysis made by M. Klein at meeting of Division of Fluid Dynamics, American Physical Society, San Diego, November, 1958. Paper E9.

the identification and characterization of

RUMBLE AND THUD

E. S. Starkman, University of California and consultant to California Research Corp.

W. E. Sytz, California Research Corp

A RECENT PROBLEM associated with combustion in high-compression-ratio spark-ignition engines is the occurrence of two noises characterized by being of much lower audio-frequency, about 600–2000 cps, than the usual 5000–8000 cps sounds of knock. While there is some confusion over usage of the terms, one of these, usually called rumble,[1] is associated with combustion-chamber deposits and is aggravated by increasing compression ratio and engine speed. Another, called thud, while sounding the same as rumble and reacting similarly to engine variables as rumble, can be brought about by advancing the spark setting.[2] The investigation reported here is concerned principally with rumble. However, thud was also observed.

While a considerable amount of information is available to indicate that rapid rates of pressure rise are associated with rumble[3,4,5] and while rumble can be produced at will in engines having compression ratios of 11/1 or more, little is available to indicate the origin of the audio-vibration or its precise character. It was the purpose of this investigation to clarify the question of source and to determine the interrelationships between rumble and engine operation. The investigation was guided by the preconceived theory that the noise of rumble was somehow related to a vibration of the crankshaft or of the power train. Therefore, the instrumentation to be utilized had to incorporate means for discerning crankshaft vibration. In addition, simultaneous pressure-time records, engine exhaust temperatures, and power output were to be obtained.

Equipment and Procedure

Instrumentation

A 1956 model V-8 engine equipped with 11/1 compression ratio heads was used for the investigation. The engine installation with its attendant instrumentation is shown in Fig. 1. Power was absorbed and measured with a Midwest eddy current dynamometer.

The principal item of instrumentation is a mag-

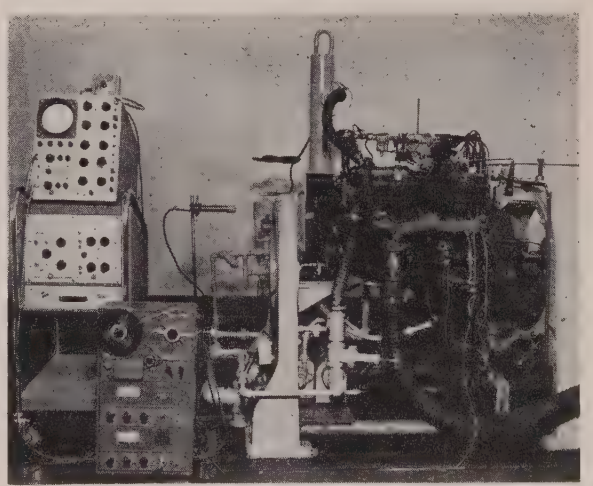


Fig. 1 — General view of test station and equipment

[1] "Terms for Use in Otto Cycle Engine Combustion." CRC Report No. 278, June, 1954.

[2] "Thudding in High-Compression-Ratio Engines," by W. E. Morris and D. C. Fariss. SAE Transactions, Vol. 67, 1959, pp. 134–137.

[3] "Rumble — Deposit Effect at High Compression Ratios," by A. E. Felt, J. A. Warren, and C. A. Hall. SAE Transactions, Vol. 67, 1959, pp. 138–144.

[4] "Engine Pounding — Its Causes and Control," by J. L. Bame and R. G. Tuell. SAE Transactions, Vol. 67, 1959, pp. 158–168.

[5] "Investigating Rumble in Single-Cylinder Engines," by J. A. Robison, M. D. Behrens, and R. G. Mosher. SAE Transactions, Vol. 67, 1959, pp. 169–174.

SIMULTANEOUS RECORDINGS of cylinder pressure, audible sound, and crankshaft motion have shown that rumble is a noise associated with bending vibrations of the crankshaft. The vibrations are caused by abnormally high rates of pressure rise near the top dead center piston position. In this study the high rates of pressure rise were obtained by inducting deposits into the the engine. Thud is a torsional vibration of the crankshaft, similar in sound to rumble but resulting from much earlier occurrence of the maximum rates of pressure rise.

Rumble vibrations consisted of a fundamental frequency of 600 cps and higher harmonics in the 11/1 compression ratio V-8 laboratory engine used in the investigation. The audible noise of rumble was predominantly composed of the second harmonic or about 1200 cps.

Limited investigation of the influence of fuel and engine operating conditions on the mechanical aspects of rumble showed no modification of the basic character or frequency of vibration but did show major effects on the amplitude and damping. Lubricating oil bulk temperature changes of as little as 52 F (162–214 F) increased the initial rate of crankshaft deflection, amplified the higher harmonics, and greatly lengthened the persistence of the vibration.

Rumble induced by dropping combustion-chamber deposits into the carburetor was accompanied by power losses as high as 19%, cooling load increases of up to 50%, and a 10% reduction in exhaust gas temperature.

After having been subjected to an inordinate amount of rumble, approximately 3 hr, the engine was disassembled and examined. No signs of mechancial damage could be detected, implying that rumble while being undesirable is apparently not destructive.

This paper received the 1959 Horning Award.*

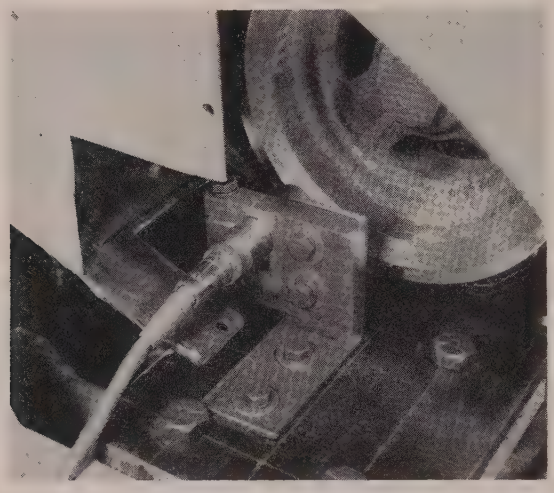


Fig. 2 — Magnetic pickup unit and crankshaft pulley

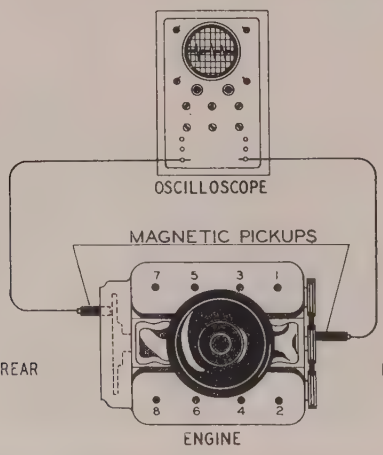


Fig. 3 — Arrangement of instrumentation

netic pickup unit for discerning abnormal crankshaft movements. The magnetic pickup unit consists of a length of 0.25-in. diameter Alnico magnetic bar stock wrapped with copper wire. One of these units is shown in Fig. 2 as it was placed in front of the crankshaft pulley and separated from it by an air gap of about 0.010 in. The wire was brought out for direct connection with no preamplification to the recording instrument, in this case an oscilloscope. The output voltage from the magnetic pickup is proportional to the time rate of change of distance between the pickup and the crankshaft pulley. A hole drilled in the pulley face also impresses a voltage transient on the system, and this was utilized to provide timing marks for the signals being recorded.

In application, the location of two of these magnetic units is shown schematically in Fig. 3. One is placed at the flywheel end of the engine and serves to trigger and synchronize the oscilloscope. The other is placed at the crankshaft pulley end and produces the signal to be interpreted. Also shown in Fig. 3 is the cylinder numbering system to be referred to later.

Fig. 4 illustrates the principle of operation of the magnetic indicator and interpretation of the oscilloscope record. In the absence of significant crankshaft displacement or vibration, the signal from the magnetic pickup unit was observed on an oscilloscope as a horizontal line interspaced with timing marks at convenient intervals of 90 deg. With torsional vibration, the distance between timing marks varies from cycle to cycle in the oscilloscope trace. This is because triggering and recording are accomplished at opposite ends of the crankshaft, and the torsional vibrations appear as differential instan-

* Paper presented at SAE Summer Meeting, Atlantic City, June 17, 1959.

taneous angular velocities of those two ends. A bending of the crankshaft appears as a vertical displacement of the signal on the oscilloscope screen; and if vibrational in character, the signal has the same frequency as the vibration.

Since it was important in this investigation that the engine be made to rumble on demand, the normal method of obtaining rumble by the buildup of deposits in the combustion chamber was not employed. Instead, rumble was obtained by metering combustion-chamber deposits directly into the intake manifold of the engine. These deposits were obtained from another engine and were not sieved or otherwise quality controlled.

In order to bring about multiple cylinder rumble, approximately one cubic centimeter of deposit was slowly aspirated into the carburetor intake. When it was desired to make only one cylinder rumble at a time, deposits were introduced directly into the intake port of the cylinder with a metering device made from a modified glass stopcock. This device is shown just to the left and in front of the distributor in Fig. 5. Modification of the stopcock consisted of plugging all but about 0.01 cc of the hole in the stopcock body, thus providing a pocket for metering deposits.

Other instrumentation noteworthy in Fig. 5 are the quartz crystal spark-plug-type pressure pickup in Cylinder No. 7 and the fine wire (No. 28 gage) chromel-alumel thermocouple in the exhaust port. At the back of the engine can be seen the magnetic pickup unit which was used for oscilloscope triggering.

The oscilloscope, a Tektronix Type 535, was equipped for dual trace so that simultaneous records of two functions could be observed and recorded. The oscilloscope also provided for the observing and photographing of a single sweep.

Engine Operating Conditions

Since the principal purpose of the investigation was to study rumble per se, the engine operating conditions were chosen as most convenient to fuel consumption, extraneous noise, and other considerations. Wide-open throttle was used for obtaining rumble, but high speeds were not found to be required. The engine was run at 1400 rpm unless otherwise specified. Spark timing under these conditions was 14 deg btdc. Oil and cooling jacket temperatures were held at about 175 F. Benzene, isooctane, and on occasion isooctane plus 6 ml tel were the only fuels used with benzene having the highest proportionate operating time.

Experimental Results

Character of Rumble

Fig. 6 is an oscillograph record of the output signal from the magnetic pickup unit while Cylinder No. 1 was being caused to rumble. The sweep was from left to right. The upper trace corresponds to 100 deg of crank angle in the compression and expansion strokes, showing a single cycle in which audible rumble of moderate level was occurring. The lower trace is the same period during a nonrumbling cycle. Engine conditions for both traces were the same: 1400 rpm, wide-open throttle, benzene fuel, spark timing 14 deg btdc, lubricating oil and cooling jacket temperatures 175 F. Fig. 6 is one of a great many of such cycles which were obtained and analyzed. They were all of this form.

The wave form of the rumble cycle shown in Fig. 6 consists of at least two and perhaps more highly damped vibrations of a frequency corresponding

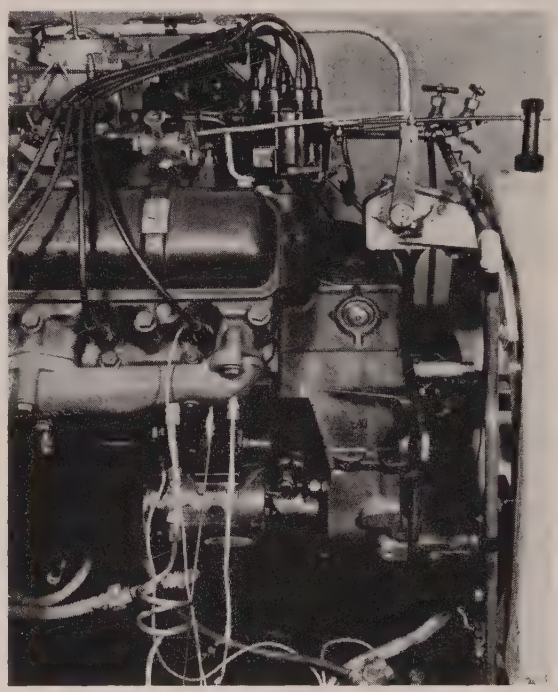


Fig. 5 — Instrumentation at No. 7 Cylinder

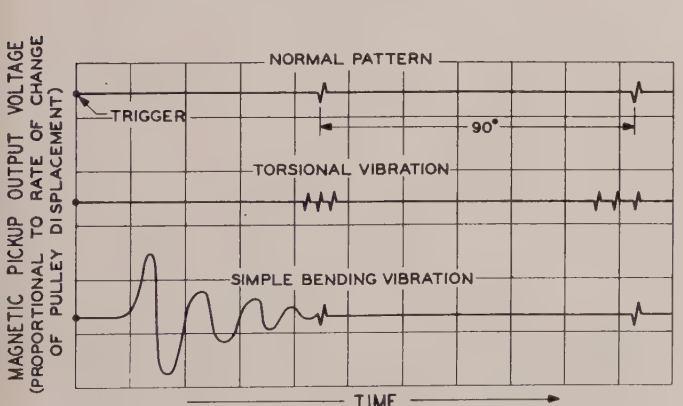


Fig. 4 — Interpretation of magnetic pickup unit oscillograph record

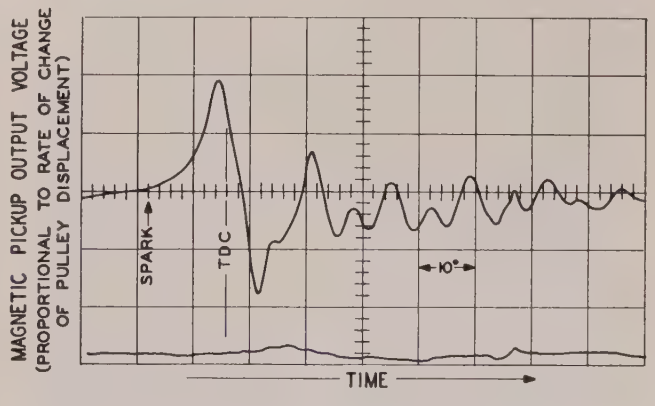


Fig. 6 — Pattern of mechanical vibration during rumble in No. 1 Cylinder (1400 rpm, wide-open throttle, benzene fuel)

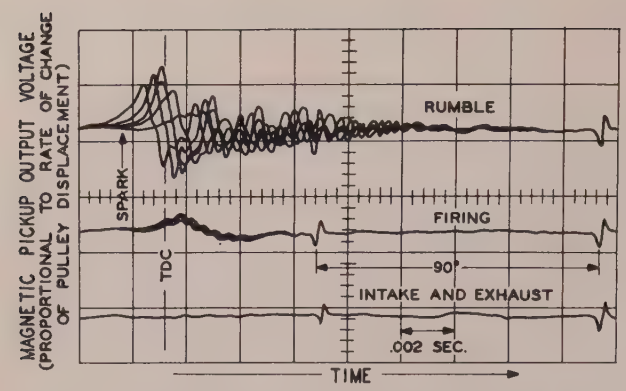


Fig. 7 — Consecutive patterns of mechanical vibration during rumble in No. 1 Cylinder (1400 rpm, wide-open throttle, benzene fuel)

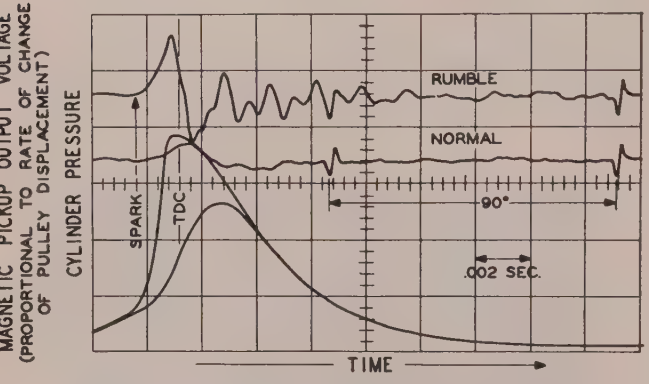


Fig. 8 — Simultaneous records of combustion pressure and mechanical vibration (1400 rpm, wide-open throttle, benzene fuel)

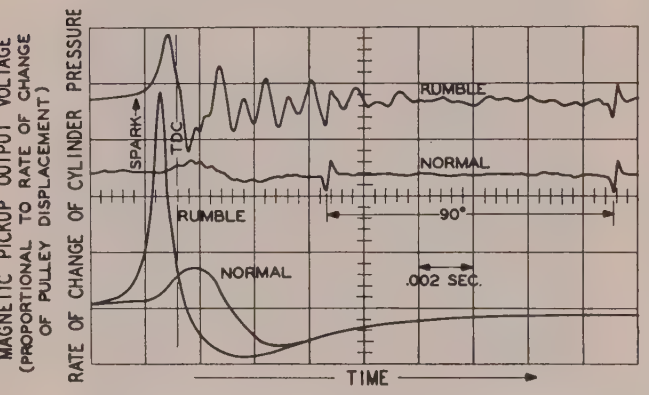


Fig. 9 — Simultaneous records of rate of pressure change and mechanical vibration (1400 rpm, wide-open throttle, benzene fuel)

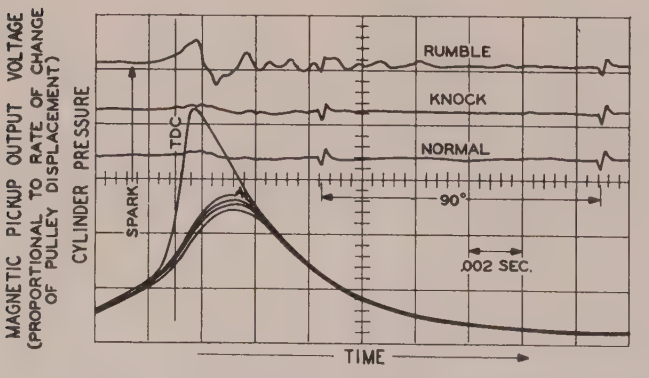


Fig. 10 — Knock, rumble, and normal combustion — combustion pressures and mechanical vibrations (1400 rpm, wide-open throttle, fuel: isooctane and isooctane + 6 ml tel)

to 610 cps and multiples thereof. The vibration almost entirely damps out in 90 deg of crank angle.

An analysis of the rumble cycle record shows that the crankshaft pulley started to move in the direction of the magnetic pickup unit about 12–15 deg ahead of tdc for Cylinder No. 1. Maximum rate of displacement occurred at about 2 deg btdc. Maximum displacement was at 3 deg after. It will be shown later that the deflections were the result of a bending of the crankshaft attendant to the high rates of pressure rise during the rumble cycle.

Fig. 7 is a record of the characteristics of the signal from the magnetic pickup unit during the intake and exhaust strokes of Cylinder No. 1 (lower trace), during a series of normal firing cycles (center trace), and during a series of six consecutive firing cycles while the engine was rumbling (upper trace). The record of the upper trace was taken during a period of rumble following the introduction into Cylinder No. 1 of a measure of combustion-chamber deposits. The six cycles which rumbled as a consequence of one deposit aspiration are of the same frequency (610 cps and multiples thereof), but the onset and amplitude of the vibrations differ considerably from cycle to cycle.

Relation between Cylinder Pressures and Rumble

Fig. 8 is a comparative record of the crankshaft vibration and of pressure in the combustion chamber during a cycle in which there was normal combustion and one in which rumble was occurring. The oscillograph trace encompasses about 166 deg of crank angle, starting at 25 deg btdc on Cylinder No. 1. The typical pattern of rumble is evident in the uppermost trace with the one below being normal combustion. The pressure history was obtained from a specially constructed quartz crystal pressure sensor built into the spark plug. Respective records of the cylinder pressure shown as the two lower curves indicate the abnormal pressures and rates of pressure rise which invariably accompany the induced crankshaft vibration. The peak pressure during normal combustion was about 600 psi; that during the rumble cycle was about 900 psi.

Fig. 9 is an illustration of the relationship between rate of change of cylinder pressure and the vibrational phenomenon identified as rumble. The upper two traces are the signals from the magnetic pickup with a normal and with a Cylinder No. 1 rumble cycle. The lower two traces are concurrent *dp/dt*, or rate of pressure change diagrams. The rate of pressure rise during rumble is five times that occurring during normal combustion. The maximum rate of pressure rise during rumble appears in this record to precede tdc by about 5 deg, and the maximum rate of crankshaft deflection occurs about 2 deg btdc.

The fuel used in the results presented to this point was benzene, which did not knock in the test engine under wide-open throttle conditions, although there were some occasions when a trace rumble was detected. While operating on isooctane the engine would give borderline knock, and on isooctane plus 6 ml tel the engine would neither knock nor rumble unless induced to rumble by the introduction of combustion-chamber deposits into the intake system. Fig. 10 is presented to show the relative influence on crankshaft vibration of rumble, knock,

and normal combustion.

As previously mentioned, the pressure and magnetic pickup records were obtained concurrently for each condition. The uppermost traces of each set in Fig. 10 are for rumble induced through the aspiration of deposits into Cylinder No. 1 using isooctane plus 6 ml tel as fuel. The middle traces are those with isooctane and show knock. This pressure trace includes the typical knock vibration shortly after tdc. The lower traces are the results obtained in the absence of knock or rumble. Fig. 10 also illustrates the observation that maximum rate of crankshaft deflection comes atdc when using isooctane, while it usually occurred btdc with benzene.

Detection and Cylinder Location of Multicylinder Rumble

The damped nature of the vibration of rumble in one cylinder only, resulting in a persistence of the order of only 90 deg of crank angle, makes it possible to identify rumble in any cylinder by observing the output from the magnetic pickup unit over 720 deg of crank angle. Illustrative of this characteristic is the oscillograph shown in Fig. 11. The top trace was made while deposits were being introduced into the carburetor. This trace shows that all cylinders except No. 6 were caused to rumble by the introduction of deposits. The middle trace was obtained with deposits introduced into Cylinder No. 1. The bottom trace was made with no deposit aspiration.

The location of the magnetic pickup at the front end of the engine and its alignment relative to the two cylinder banks results in differences in vibrational patterns for the various cylinders. The most apparent difference is the decrease in amplitude of the vibration with increasing distances of cylinders from the magnetic pickup. This is primarily due to damping.

Sound of Rumble

A high fidelity microphone was connected to the oscilloscope at the same time as the magnetic indicator. The microphone was held about 1 in. from the left rocker box cover while Cylinder No. 1 was rumbled. The outputs recorded from the two units simultaneously are shown in Fig. 12. The characteristic appearances of normal and rumble cycles are apparent in the two top traces. The patterns of the acoustical phenomena were simultaneously recorded for the two conditions as shown.

The acoustical vibration had a frequency of about 1200 cps, which corresponds within normally expected accuracy to the second harmonic of the magnetic pickup trace. The time displacement between the start of the magnetic and acoustic vibrations reflects the time necessary for the mechanical vibration to excite the metal and air gap separating the source of the noise and the microphone which was used to record the phenomenon.

Nature of the Crankshaft Oscillation with Rumble

In the early phases of the investigation, the pattern of rumble as recorded from the single magnetic unit at the pulley was identified as being related to an induced periodic motion of the crankshaft. It was in a lateral plane and occurred at a frequency of about 600 cps and multiples thereof. The voltage output records from the instrumentation were interpreted as indicating that the crankshaft pulley initially moved away from the engine. After this first deflection, a highly damped periodic oscillation took place.

It was suspected that this initial deflection was one of bending, and the oscillation was a manifestation of crankshaft lateral or bending vibrations.[6] To establish the validity of this theory, a universal

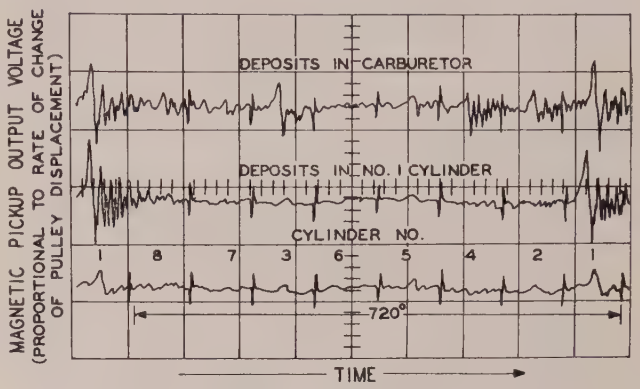

MAGNETIC PICKUP OUTPUT VOLTAGE (PROPORTIONAL TO RATE OF CHANGE OF PULLEY DISPLACEMENT)
DEPOSITS IN CARBURETOR
DEPOSITS IN NO. 1 CYLINDER
CYLINDER NO.
1 8 7 3 6 5 4 2 1
720°
TIME

Fig. 11 — Mechanical vibrations of crankshaft during multicylinder rumble (1400 rpm, wide-open throttle, benzene fuel)

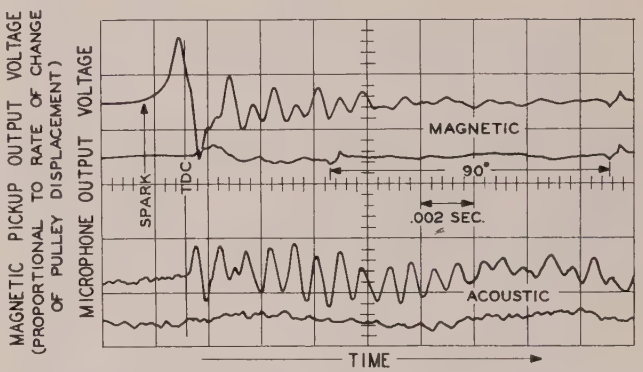


Fig. 12 — Simultaneous sound and mechanical vibrations with rumble (1400 rpm, wide-open throttle, benzene fuel)

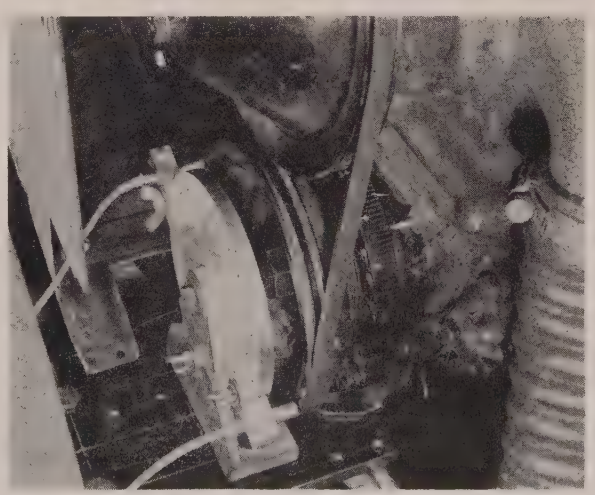


Fig. 13 — Dual unit universal magnetic indicator assembly

[6] "Crankshaft Vibration," by W. K. Wilson. *Gas and Oil Power*, Vol. 53, April, 1958, p. 98.

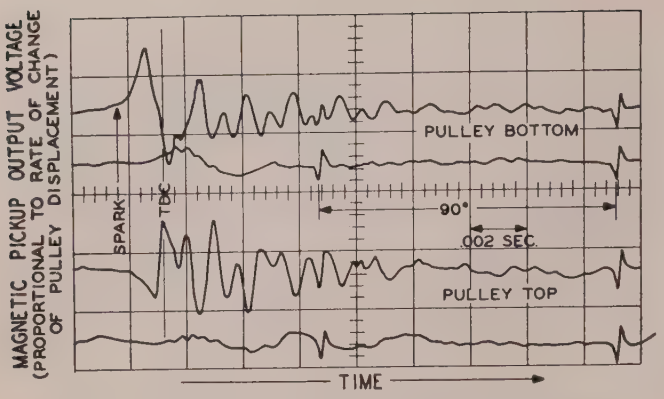


Fig. 14 — Nature of crankshaft vibration during rumble in No. 1 Cylinder (1400 rpm, wide-open throttle, benzene fuel)

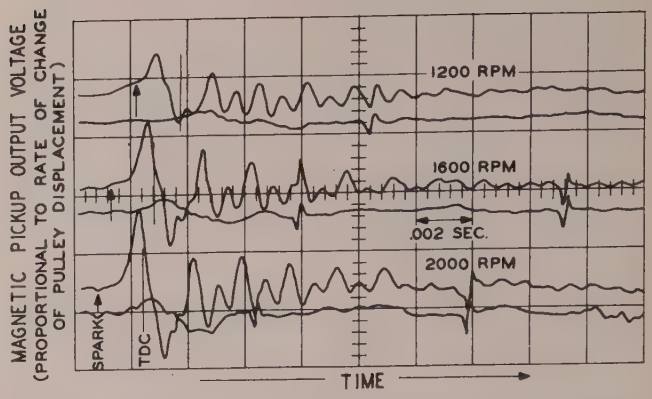


Fig. 16 — Influence of engine speed on crankshaft vibration during rumble in No. 1 Cylinder (wide-open throttle, benzene fuel)

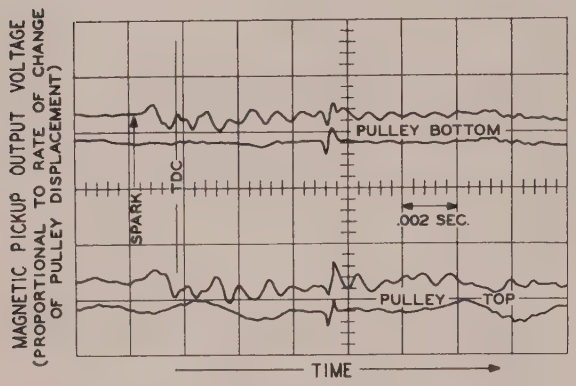


Fig. 15 — Nature of crankshaft vibration during rumble in No. 7 Cylinder (1400 rpm, wide-open throttle, benzene fuel)

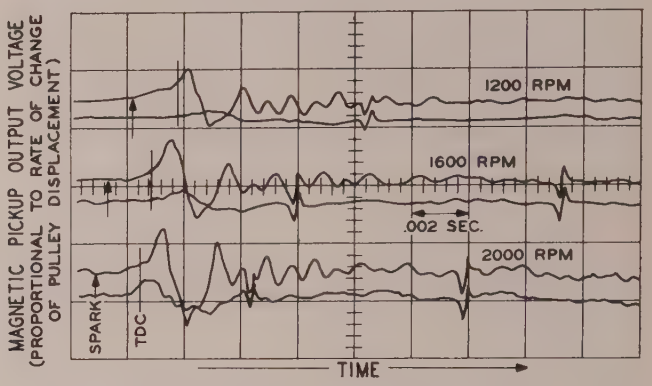


Fig. 17 — Influence of engine speed on crankshaft vibration during rumble in No. 1 Cylinder (wide-open throttle, isooctane fuel)

holder was built for positioning two magnetic indicators in front of the crankshaft pulley in any desired radial location. With this scheme it was then possible to record simultaneously the movement of the pulley in the lateral direction at two points. The installation is shown in Fig. 13. The magnetic pickup units used were $\frac{3}{8}$ in. outside diameter and $1\frac{1}{2}$ in. long. One-eighth in. diameter Alnico magnet bar stock was used.

Fig. 14 is an oscillogram of the simultaneous voltages from the two units when placed 180 deg apart in the plane of Cylinder No. 1 while it was being rumbled and while it was undergoing a normal firing stroke. It is readily apparent that during rumble the crankshaft pulley first was deflected in a manner which could only be due to crankshaft bending and that this was followed by an oscillation in a bending mode. The bending is indicated by the bottom of the pulley always moving in opposite direction to the top.

Cylinder No. 7 was also rumbled and the instrumentation aligned in the same manner as for Cylinder No. 1. Fig. 15 shows the results. The amplification and time scale were the same for Figs. 14 and 15. The reduced amplitude of the vibration from Cylinder No. 7 as compared to Cylinder No. 1 might be due to damping.

The record obtained at the crankshaft pulley with Cylinder No. 7 rumbling is in its initial phases dissimilar to that of Cylinder No. 1; the movement commences almost entirely as a lateral or axial displacement of the end of the crankshaft away from the front of the engine. However, after two oscillations the pulley action is similar to that with Cylinder No. 1. The interpretation is that the initial deflection of the crankshaft at Cylinder No. 7 is transmitted as an axial vibration of the shaft, the bending oscillation follows but is greatly damped.

Influence of Selected Engine Variables on Rumble

Influence of Speed — The engine was operated at fixed conditions as in previously presented data, but speed was varied over a selected range. Fig. 16 shows the results at 1200, 1600, and 2000 rpm using benzene as the fuel. Spark timing varied during these tests as dictated by the normal distributor advance. They were, respectively, 10, 16, and 22 deg btdc for the three speeds. The time scale has been held constant in the oscillogram. For each speed a normal and a rumbling cycle were recorded from Cylinder No. 1.

Fig. 16 makes it graphically obvious that except for amplitude the vibration associated with rumble is substantially independent of engine speed. The 610-cycle fundamental frequency of the oscillation is the same, as are the characteristics of the harmonics from which the oscillations are composed and to some extent the time necessary for them to damp out. The amplitude of the vibration increased with speed.

Influence of Fuel Composition — A previous section of this report dealt with the relative rumble and knock characteristics of benzene, isooctane, and

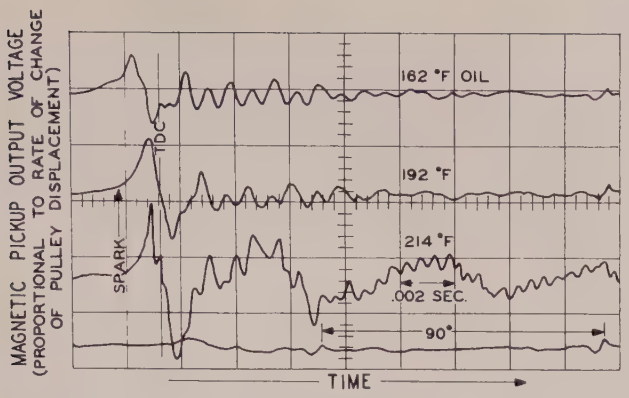


Fig. 18 — Effect of lube oil bulk temperature on crankshaft vibration during rumble in No. 1 Cylinder (1400 rpm, wide-open throttle, benzene fuel)

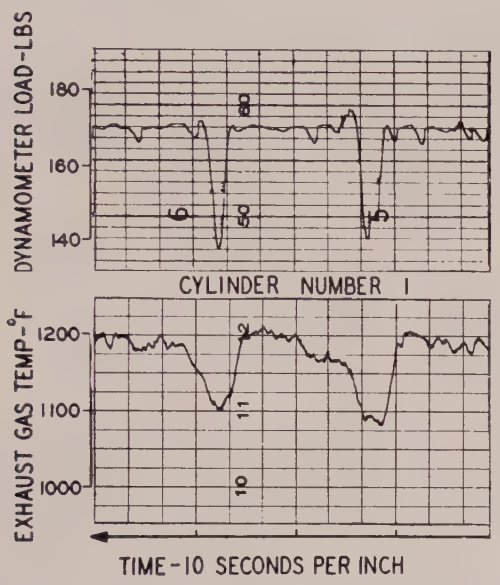


Fig. 19 — Influence of rumble on power and exhaust temperature with deposits inducted into carburetor (11/1 compression ratio, 1400 rpm, wide-open throttle, benzene fuel)

isooctane plus 6 ml tel. As was previously known[3,4] and as was recognized during this investigation, rumble is more easily obtained and is more pronounced with benzene than with isooctane. The presence of as much as 6 ml tel in the isooctane had little effect on rumble. To illustrate further the influence of fuel character on the mechanical aspects of rumble, the tests at variable speeds were repeated using isooctane as fuel. Results are shown in Fig. 17 at 1200, 1600, and 2000 rpm. The electronic amplification of the signals was the same as with benzene. It can be seen that the amplitude of vibration has decreased, the location of maximum deflection has moved to a position atdc, and the length of time for the vibration to damp out has been reduced. The frequencies of vibration are the same.

Influence of Bulk Oil Temperatures — Tests were made with crankcase oil temperature controlled to 162, 192, and 214 F with other conditions held constant. The results are shown in Fig. 18. With the bulk oil temperature ranging from 162 to 214 F, it is apparent that the vibration of rumble changes radically in character. The item of major note is the increased rate of first deflection and the increasing prominence of the third harmonic. This third harmonic is evident in the record with the 214 F bulk oil temperature. Also evident in this same record is a relatively low frequency oscillation which would appear to correspond to one-third to one-fourth the fundamental frequency of 600 cps.

Character of Thud

An intermittent noise similar to rumble occurred with increasing frequency after about 200 hr of testing. The engine was run mostly at full throttle. The noise was not associated with the induction of deposits into the engine but appeared to be voluntarily generated within the engine. It was not accompanied by the usual oscillations of the crankshaft related to rumble. The rumblelike noise grew progressively worse, and finally the engine backfired through the carburetor.

After deliberation, it was deduced that the phenomenon was linked to the combustion process and was possibly preignition. The pattern of noise and backfire was thereafter reproduced by manually overadvancing the distributor timing. It was noted during this period that the timing marks on the oscillograph were being displaced horizontally, an indication of torsional deflection and, therefore, probable torsional vibration of the crankshaft. Unfortunately, the phenomenon was not photographed at the time; but it was subsequently recognized as being what is described as thud since the phenomenon and its method of production closely corresponded to that described in a recent SAE publication.[2]

The occurrence of thud was eliminated by replacing the spark plugs with new ones of a lower heat range. The original set of plugs appeared to have been operating very hot. The plug in Cylinder No. 1 when examined under a microscope showed slight surface damage to the ceramic. This damage was probably due to the fact that Cylinder No. 1 had undergone considerable induced rumble. It is also likely that it was Cylinder No. 1 which was observed thudding since the oscilloscope pattern was indicating highly variable angular velocities during the working stroke in Cylinder No. 1.

Effect of Rumble on Engine Performance

There was reason to believe that the high rates of pressure rise and the shift in peak pressures accompanying rumble should have some influence on the energy distribution in the engine. For this reason, a brief investigation was made of the changes which might occur. The investigation was designed for rough determinations only, and no measurements were made of the flow rates of air or water. Data obtained consisted of fuel flow rate, brake horsepower output, and exhaust gas temperature from one cylinder. A more precise determination of engine thermodynamic functions was considered beyond the scope of the investigation.

Fig. 19 is typical of the exhaust temperature and power output records obtained with deposits being aspirated into the carburetor to induce multicylinder rumble. The fuel was benzene. An examination of the upper record of engine output shows that in the absence of rumble there was about 170-lb force on the dynamometer lever arm. This corresponds to 79.5 hp and represents 32% of the fuel

energy. When rumble was induced at a level which would be termed medium-heavy (the two large inverted humps in the record), the output fell off by 32 lb or 19%. This drop in power output was accompanied by a reduction in exhaust gas temperature as shown in the lower record for Cylinder No. 1. While the temperatures indicated are far from being accurate, it can be noted that a drop of 110 F out of 1200 F accompanied rumble in Cylinder No. 1. This represents about a 10% reduction of the total amount of energy being rejected to the exhaust. The exhaust gases were estimated to carry away about 35% of the heating value of the fuel in these tests.

A simple accounting for the energy being supplied to the engine, taking cognizance of the reduction in power and exhaust temperature, leads to the conclusion that the load on the cooling system must have been proportionately increased.[7] For the conditions of the test shown in Fig. 19, about 20% of the fuel energy goes to the cooling jacket when rumble is absent. With rumble present to the extent of medium-heavy, the cooling load must have been increased by about 10% of the fuel energy to account for the power loss and drop in exhaust temperature. This is an increase of about 50% in load on the cooling system.

Effect of Rumble on Engine Mechanical Condition

The engine was operated approximately 300 hr during the investigation reported here. It is difficult to make an exact determination of how much of this time the engine was rumbling in one or more cylinders. It was estimated to have been about 2½ hr of single-cylinder rumble and ½ hr of multicylinder rumble. The extent to which the engine was rumbled was well beyond that which might be expected in the normal lifetime of an engine.

The engine was disassembled and all mechanical parts examined. The crankshaft and pistons were subjected to inspection for cracks by Magnaflux and Zyglo, respectively, and found free of any such defects. There was complete absence of evidence that rumble had damaged or worn any of the other parts, including main and rod bearings, pistons, valves, rings, and cylinders. It was also noted that the piston crown and combustion chamber of Cylinder No. 1, which had been subjected to the major amount of rumble, were coated with a light deposit which appeared impacted into place. Rumble may also have decreased spark-plug life, but further tests would be required to establish this.

Discussion

The experimental evidence of the investigation makes it apparent that the engine noise which is called rumble is composed of the same frequencies as the measured mechanical vibration of the engine crankshaft in a bending mode. The bending must of necessity also include an axial component.

Results from other laboratories which describe rumble and thud as having a sound spectrum which extends from 400 to 2000 cps[5] are easily qualified. There is a tendency for the crankshaft to vibrate in a complex form of its various natural frequencies. The broad spectrum reported by other investigators, reflects the lack of definition inherent in the equipment used for sound spectrum analysis.

The induced vibrations of the crankshaft are definitely linked to abnormally high rates of pressure rise. When these high rates of pressure rise are far ahead of tdc and could be termed preignition, there is sufficient instantaneous reversed torque to induce a restraint on the crankshaft. This gives rise to a torsional vibration and so-called thudding. Advancing the spark is one way to induce this phenomenon.

When the high rates of pressure rise occur near tdc, the crankshaft receives a bending impulse. This results in rumble. There is no reason to believe that one phenomenon cannot blend into the other, and furthermore a coupled vibration of bending and torsion is highly possible.

Both fuel composition and lubricating oil temperature were shown to influence the intensity of the mechanical manifestation of rumble. Fuel composition studies were extremely limited consisting only of comparisons between benzene, isooctane, and isooctane plus 6 ml tel. Further studies are needed to clarify the influence of fuel and lubricant properties and other factors on the incidence and intensity of the phenomenon.

Conclusions

1. A method has been developed to identify and characterize rumble and thud.
2. Rumble is a bending vibration of the crankshaft.
3. Thud is a torsional vibration of the crankshaft.
4. Rumble is induced by high rates of combustion pressure rise near the tdc position of the piston; thud is due to high rates of pressure rise occurring earlier in the stroke.
5. The major audible component of rumble is the second harmonic of the mechanical vibration or about 1200 cps.
6. The onset, amplitude of vibration, and the relative contributions of the various harmonics of rumble are influenced by fuel and engine operating factors. However, these items have no influence on the frequencies of the individual components of the vibration.
7. Under wide-open throttle conditions, a power loss of 19% was observed to accompany rumble. Exhaust gas energy fell 10% and cooling system load increased 50%.
8. Each induced vibration of rumble damps out in about 90 deg of crank angle. The phenomenon is repititious; each successive cycle has the same appearance as that preceding and following.
9. The engine although subjected to deliberate and extensive rumble did not suffer mechanical damage.
10. Spark-plug life was decreased by rumble in these tests.

Acknowledgments

This work represents the combined efforts of many people among whom J. H. Macpherson, K. L. Kipp, and V. C. Davis must be particularly singled out for their contributions.

[7] "Internal Combustion Engines," by L. C. Lichty. Pub. by McGraw-Hill Book Co., New York, 1951.

ROLLS-ROYCE DIESEL ENGINES

R. H. Whiteside
Rolls-Royce of Canada Ltd.

ROLLS-ROYCE diesel engines were designed so that they would be adaptable for many different applications. From a simple engine design, based on a cylinder of 5⅛-in. bore and 6-in. stroke, engines were developed in the following basic arrangements:

- Four, six, and eight cylinders, with swept volumes of 495, 743, and 990 cu in., respectively.
- Turbocharged, supercharged, and normally aspirated.
- Wet- or dry-sump lubrication.
- Left- or right-hand build.
- Horizontal or vertical configuration.
- Single or coupled units.
- A variety of take-off points, with full engine power from the front of the crankshaft.

This paper describes in detail one typical engine — the supercharged dry-sump 6-cyl configuration.*

ROLLS-ROYCE diesel engines have been designed and developed since World War II. Up to that time, British development of diesel engines for automotive purposes had been mainly influenced by legal restrictions on the size and speed of load-carrying vehicles, plus the fundamental economic necessity for minimum fuel consumption. The result was that Britain concentrated on diesel power units up to about 120 hp with the 4-stroke direct-injection type of engine predominating. The post-war demand for the export of high-speed diesel engines to countries, particularly in the dollar area, for such applications as road transport, earthmoving equipment, rail traction, industrial power packs, found Britain having to design at short notice larger engines comparable to current North American practice.

Rolls-Royce designed and developed a range of pressure-charged engines consisting of three basic types: 4, 6, and 8 cyl. As each of these engines has the same bore and stroke (5⅛ × 6 in.), the swept volumes are 495, 743, and 990 cu in. respectively, the power range covered by these engines being from 90 to 600 bhp. (See Figs. 17 and 18, p. 110.)

A feature of the basic design was the introduction of the maximum number of interchangeable components which results in both economy in production and economy for the user. The in-line arrangement was chosen to permit the maximum rationalization of parts and simplified installation. The basic engine design is simple and orthodox in its conception. Complicated mechanisms have been ruthlessly eliminated. Compromise to satisfy all

* Paper presented at meeting of SAE Philadelphia Section, Feb. 11, 1959.

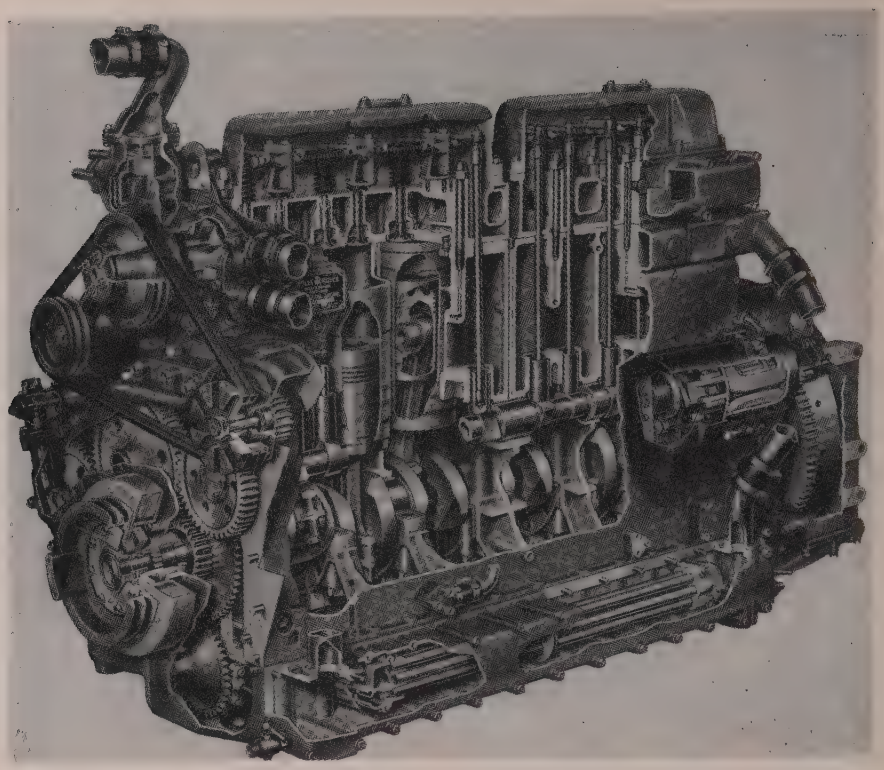


Fig. 1 — Cutaway view of Rolls-Royce supercharged 6-cyl diesel engine

the demands of such widely different duties and applications has been avoided by the use of basic engines which can be assembled in a wide variety of configurations. The versatility that has been achieved is illustrated by the following basic arrangements:

1. Four, six, and eight cylinders.
2. Turbocharged, supercharged, and normally aspirated.
3. Wet- or dry-sump lubrication.
4. Left- or right-hand build.
5. Horizontal or vertical cylinder configuration.
6. Single or coupled units.
7. A variety of power take-off points which include full engine power from the front of the crankshaft.

In order to examine some of the design features of these engines, I will deal with one typical engine in the range — the supercharged dry-sump 6-cyl engine.

Crankcase, Crankshaft, and Bearings

The crankcase has been designed to achieve strength and rigidity without sacrificing lightness. It is of monobloc construction and produced in a high-grade nickel-chrome alloy iron. Both ends of the crankcase are symmetrical which enables the engine to be assembled as a "mirror image." In order to ensure absolute freedom from contamination of the lubricating oil by residual casting sand, the crankcase is coated (prior to machining) with a phenolic resin to seal the interior surfaces. As a further precaution, all internal oilways are bored to ensure absolute cleanliness and minimum resistance to flow. (See Fig. 1.)

The main bearings are supported by deep webs and the more highly stressed center and end bearings are strengthened by the use of transverse cap screws. This means that great rigidity is obtained in the lower half of the crankcase and deflection of the bearing housings is minimized.

The crankshaft is supported by seven main bearings of the precision shell type which are prefinished and readily replaceable. The bearing comprises a steel back to which is bonded a copper lead bearing surface. To provide added strength to the bearing, assist breaking in, and prevent corrosion, the copper lead is overlaid with a thin layer of lead-indium alloy.

An unusual feature of the engine is the crankshaft which is a one-piece forging in chrome-molybdenum steel, machined all over, and balanced by removing material from the webs. After grinding, the main journals and crankpins are hardened by the nitriding process, then lapped to final size. This process has been employed for many years by Rolls-Royce on their piston engine crankshafts and, besides producing a surface having great resistance to wear, it is also a characteristic of nitriding that it produces a compressive stress in the skin. This results in added fatigue strength in the shaft.

Journals and crankpins have carefully blended radii and fine polished surfaces. Since they overlap, the crankshaft has great rigidity. Sludge traps are provided in the crankpins. The crankshaft carries at the free end of viscose silicone damper which effectively reduces the amplitude of the major critical (sixth order) from ± 0.5 to ± 0.075 deg. The natural frequency of the system is approximately 10,000 cpm.

Pistons, Connecting Rods, and Cylinder Liners

The pistons are die cast in a Rolls-Royce high silicone aluminum alloy and are tin plated after cam grinding. The combustion chamber, in the form of a toroidal cavity, is machined in the center

of the piston crown and the volume of this cavity controls the nominal compression ratio — 16/1 on normally aspirated engines and 14/1 on pressure-charged engines. (See Fig. 2.)

There are three compression and one oil control rings, all of which are mounted above the piston pin, the latter being of the fully floating type retained by snap rings. The two lower compression rings which are 1-deg taper faced, are tin plated in order to provide a good surface when breaking in. The top compression ring, however, is chromium plated and is, for certain duties, mounted in an Austenitic iron insert. The gas ring grooves are machined at a very slight angle to ensure that the ring faces have at all times a "positive rake" on the downward piston stroke.

The pistons work in centri-cast iron wet liners secured in the crankcase by flanges at their upper ends and located and sealed at their lower ends by two synthetic rubber O-rings. The liners are a slip fit and are easily fiitted and removed. (See Fig. 3.)

The cylinder liner is hardened and tempered by a patented process. This results in a ductile zone in the highly stressed area at the change of section below the flange. This heat-treatment process ensures that the iron in this zone is in the pearlitic condition, whereas the liner barrel has a martensitic structure with a high resistance to wear. The liner bores are finely ground and lapped to a surface finish of 20/45 microin.

The liner and piston combination has been carefully selected to ensure prolonged life with an extremely low wear rate, and in order to reduce ring and bore wear still further, the engine lubrication system has been designed to provide positive lubrication of the cylinder walls themselves from the very instant of start-up when the rate of wear is normally at its highest. This is achieved by means of two oil sprays directed onto each cylinder wall through drillings in the connecting rod.

The connecting rods are "H" beam, chrome-molybdenum steel forgings, drilled longitudinally to provide lubricant to the piston pin bearing and the cylinder wall spray jets previously mentioned. The shanks of the connecting rods are polished and blended to remove stress raisers. Additionally, to avoid a high stress concentration at the blend of the connecting rod shank and the big end, the bearing cap bolt heads are provided with outside locks. (See Fig. 4.)

Cylinder Head

The engine has two cylinder heads, these being identical castings in a high-grade allow iron. Each head is secured to the crankcase by 23-9/16-in. studs. These are so arranged that even loading between the cylinder head and liner flanges is obtained by the eight studs equally spaced around and close to each cylinder bore.

Complete sealing of the combustion chambers is ensured by the use of corrugated steel gaskets which are positively located by dowels.

As a further precaution to ensure a positive seal between the head and the crankcase, coolant is transferred from one to the other by means of brass transfer bobbins having O-ring seals at each end.

Replaceable shrink fit chromium iron inserts are provided for both the inlet and exhaust valves.

The injectors pass diagonally through the head in

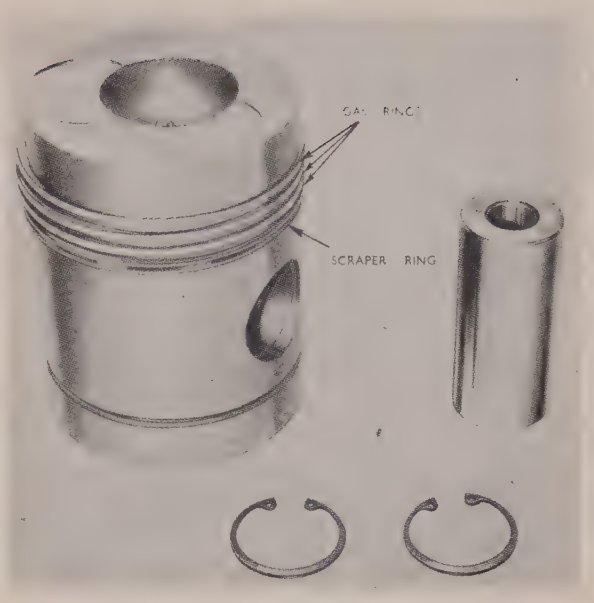


Fig. 2 — Piston and pin assembly

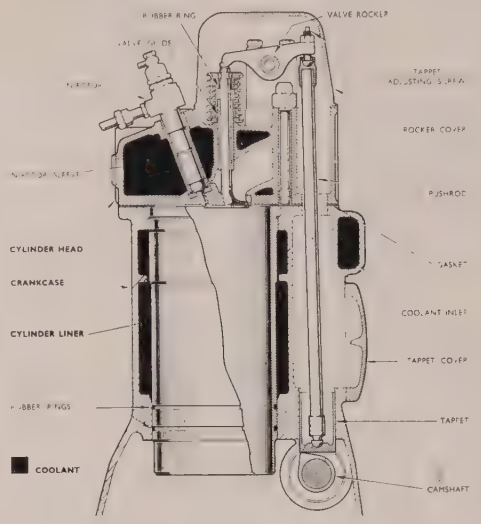


Fig. 3 — Section through cylinder

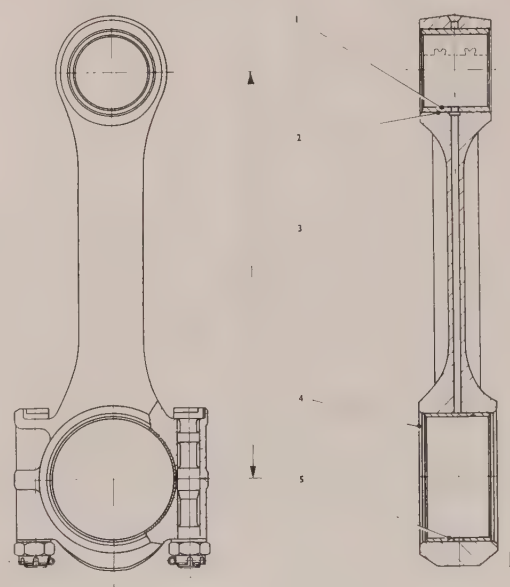


Fig. 4 — Connecting rod

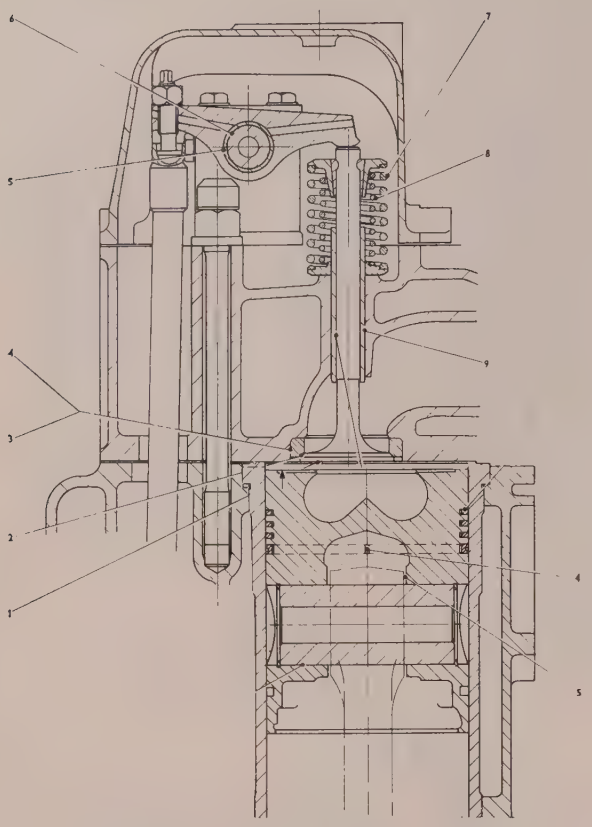
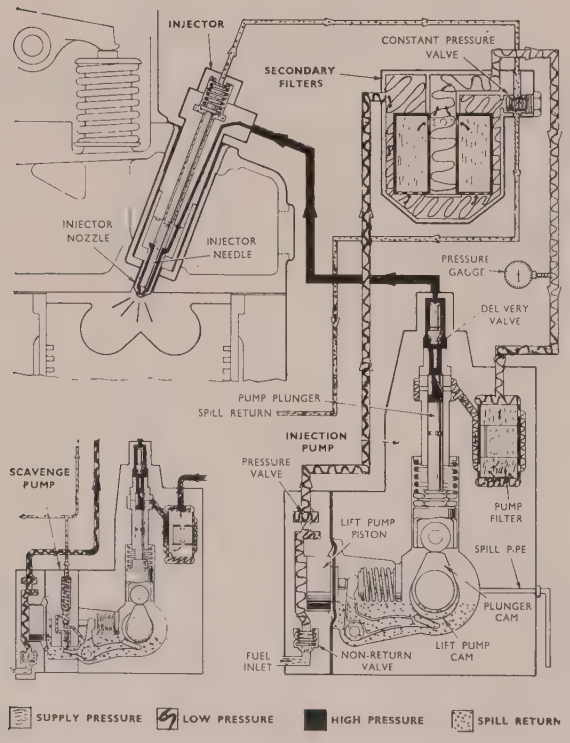


Fig. 5 — Part section through cylinder and valve gear



copper alloy sleeves which are in contact with the coolant. It will be seen that the securing studs for the injectors are outside the rocker covers and are easily accessible. (See Fig. 3.)

Valves and Valve Gear

Only two valves per cylinder are used with a consequential simplification of the valve operating mechanism, which is further simplified by the absence of injector actuators.

Both the exhaust and inlet valves are made of high-grade Austenitic steel, stellited on their faces and their tips. On pressure-charged engines, the exhaust valves are sodium filled to allow greater cooling of the valve head. Each valve has two helical springs having the same direction of helix in order to promote valve rotation. As the most common cause of valve spring failure is corrosion fatigue, the valve springs are coated with a special flexible enamel. Each valve top washer embodies a synthetic rubber oil sealing ring.

The valves are operated by rocker arms actuated by hollow push rods and plain cylindrical chilled cast-iron tappets. The tappets are offset to the camshaft lobes to ensure rotation.

The case-hardened camshaft is fully supported in seven bronze alloy bushes and these are lubricated through a rifle drilling in the shaft.

The valve gear is simple and robust and as the number of wearing parts is reduced to a minimum, correct valve lash is maintained over long periods of operation.

Combustion System

At this stage it might be appropriate to describe briefly the combustion system (Fig. 5). Towards the top of the compression stroke a high degree of "squish" is obtained by the small clearance which exists between the piston crown and the cylinder head. The air is, therefore, directed with a radial component into the toroidal combustion cavity. The shape of the cavity induces a second component with a two plane rotational swirl. At 28 deg btdc, injection of the fuel commences through a nozzle tip having four orifices. Good control of air movement over a wide speed range has been achieved without resorting to the use of masked inlet valves. That high thermal efficiency has been

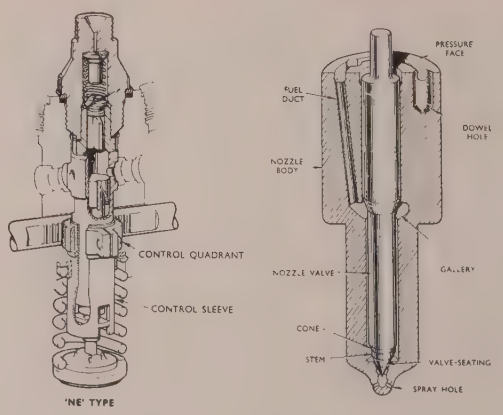


Fig. 6 — Diagrammatic view of fuel injection system

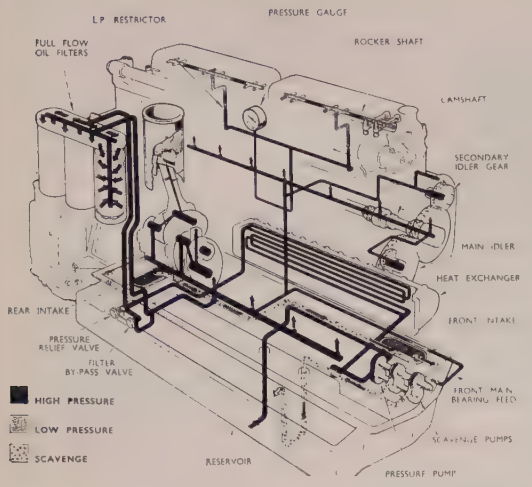


Fig. 7A — Vertical engine lubrication circuits (dry sump)

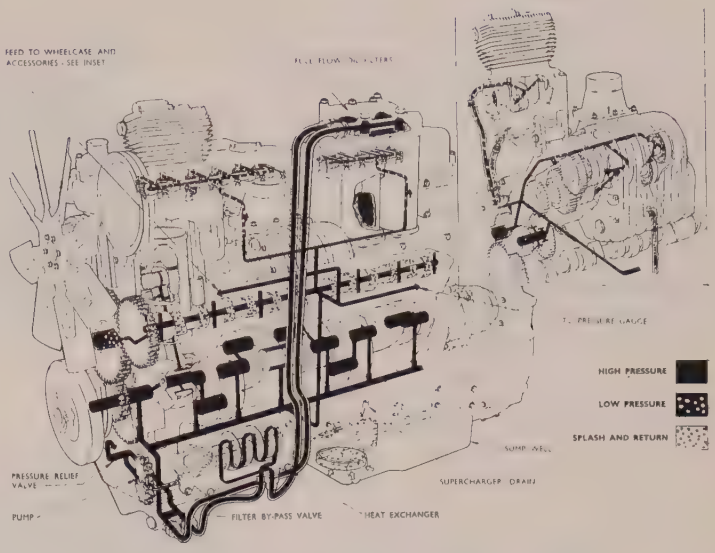


Fig. 7B — Vertical engine lubrication circuits (wet sump)

achieved with this design is indicated by the low specific fuel consumption.

Fuel Injection System

I will not spend a great deal of time elaborating upon the fuel injection system as this is basically similar to American Bosch, with which most of you are familiar (Fig. 6). The fuel injection pump can be fitted with a hydraulic or a fly-ball type governor.

As fuel cleanliness is of primary importance in ensuring long life from the injection equipment, fuel filtration has received particular attention. Fuel is subjected to 3-stage filtration in the following manner. A primary filter, usually of the paper or felt element type, is installed in the fuel line between the supply tank and the low-pressure feed pump. The fuel then flows through three more filters arranged in parallel. These contain renewable filer cartridges of paper element spiral construction which provides an extremely high degree of filtration efficiency. Since there is always the possibility that, when filter elements are renewed, traces of dirt may be allowed to pass beyond the second-stage filters, a third stage of filtration is provided to make sure that these particles are prevented from entering the pumping elements. This third stage comprises three felt elements which are incorporated inside the fuel injection pump and intentionally inaccessible without dismantling the pump itself. It is not intended that these elements be disturbed during the service life of the engine.

Lubricating Oil System

Two basic types of lubricating system are provided — either wet or dry sump (Fig. 7). Which system is utilized will depend upon the operating angles to which the engine is likely to be subjected in service.

Dry-Sump System — A separate oil reservoir is provided on the side and integral with the main oil pan. Oil is drawn from this reservoir by the main pressure pump (which is gear driven from the front end of the crankshaft) and delivered to the full-

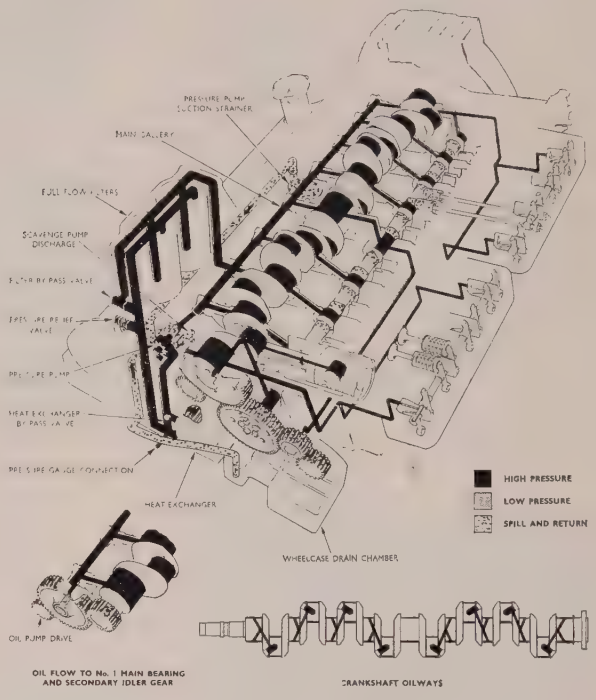


Fig. 8 — Horizontal engine lubrication circuit

flow lubricating oil filters. Two oil pressure relief valves are embodied in this part of the circuit to limit the maximum oil pressure and to provide a bypass system should the filter elements be allowed to become choked. From the three full-flow filters oil is delivered via an oil-to-coolant heat exchanger (which is also integral with the oil pan) to the main oil gallery and to the various bearing surfaces in the engine. Oil which drains back into the oil pan is then removed by two scavenge pumps, driven coaxially with the pressure pump. One scavenge pump draws oil from the forward end of the pan and the other from the rear. Oil is returned in this manner through a deaeration screen into the main oil reservoir.

Wet-Sump System — On this system oil is drawn

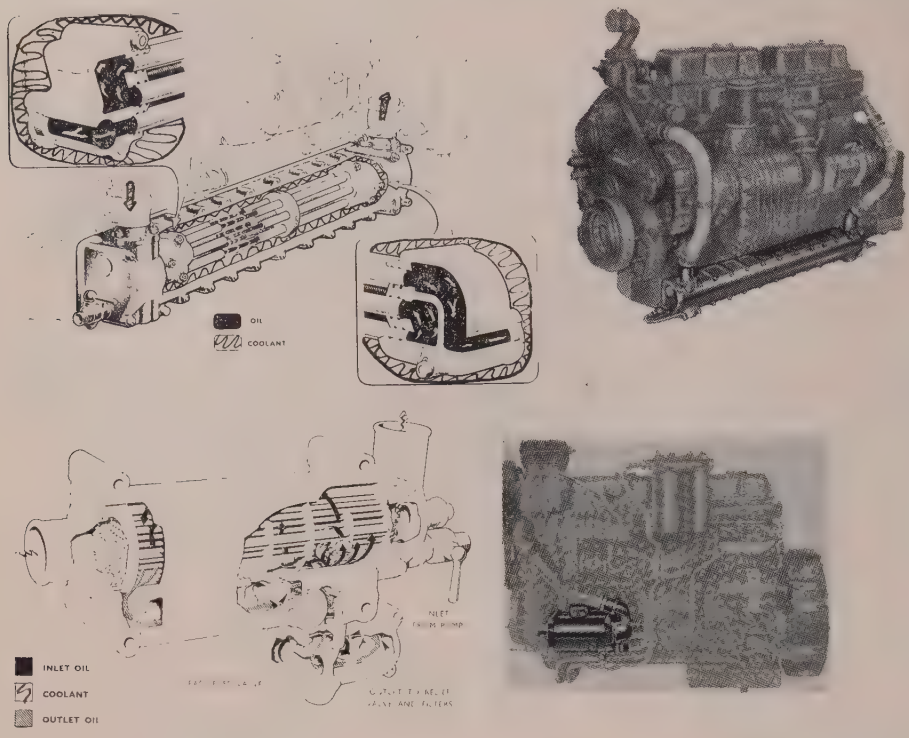


Fig. 9 — Oil-to-coolant heat exchangers

from the oil pan by a single pressure pump through a similar circuit as described above, except that in this latter case the oil-to-coolant heat exchanger is a separate pack-type unit and not an integral part of the oil pan. Oil having been fed to the bearing surfaces in the engine drains back into the pan.

Previous mention has been made of the provision of sludge traps in the crankshaft throws. These cavities are machined in the crankpins and oil is delivered to the big end bearing at a point of minimum radius. This results in solids in the oil being centrifuged and contained within the sludge traps. (See Fig. 8.)

The camshaft bearings and timing gears are lubricated by oil through a central drilling in the camshaft itself, while the rocker gear, supercharger, and the rest are fed at reduced pressure. (See Fig. 9.)

Timing Gears and Auxiliary Drives

For one "hand" of build, four different timing cases are available, all of which are common to the 4-, 6-, and 8-cyl models. These assemblies enable engines to be built either vertical or horizontally in any of the following forms: normally aspirated, turbocharged, or supercharged. Additionally, in the case of the vertical configuration the supercharger can be mounted in two alternative positions. (See Fig. 10.)

Timing and auxiliary drive gears are produced from a 3% nickel/chrome/vanadium steel, hardened and ground. Straight spur gears of 25-deg pressure angle and 8 dp are used, all gears being interchangeable as spares and do not require matching.

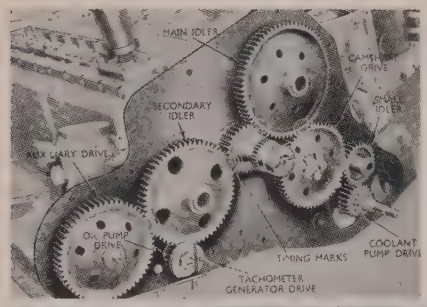
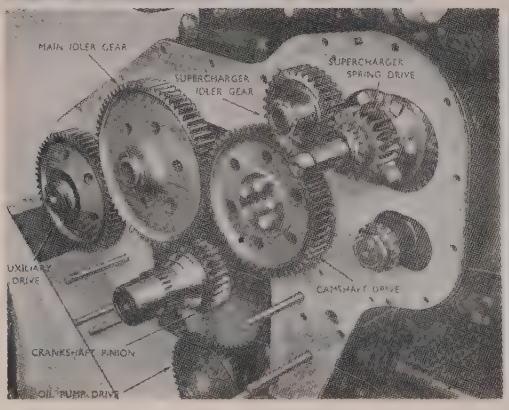
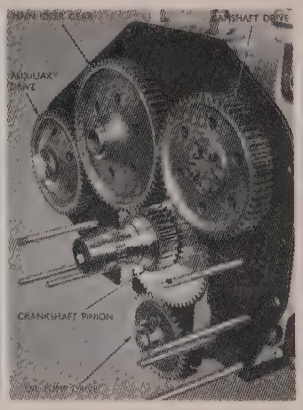


Fig. 10 — Optional timing and auxiliary gear drive arrangements





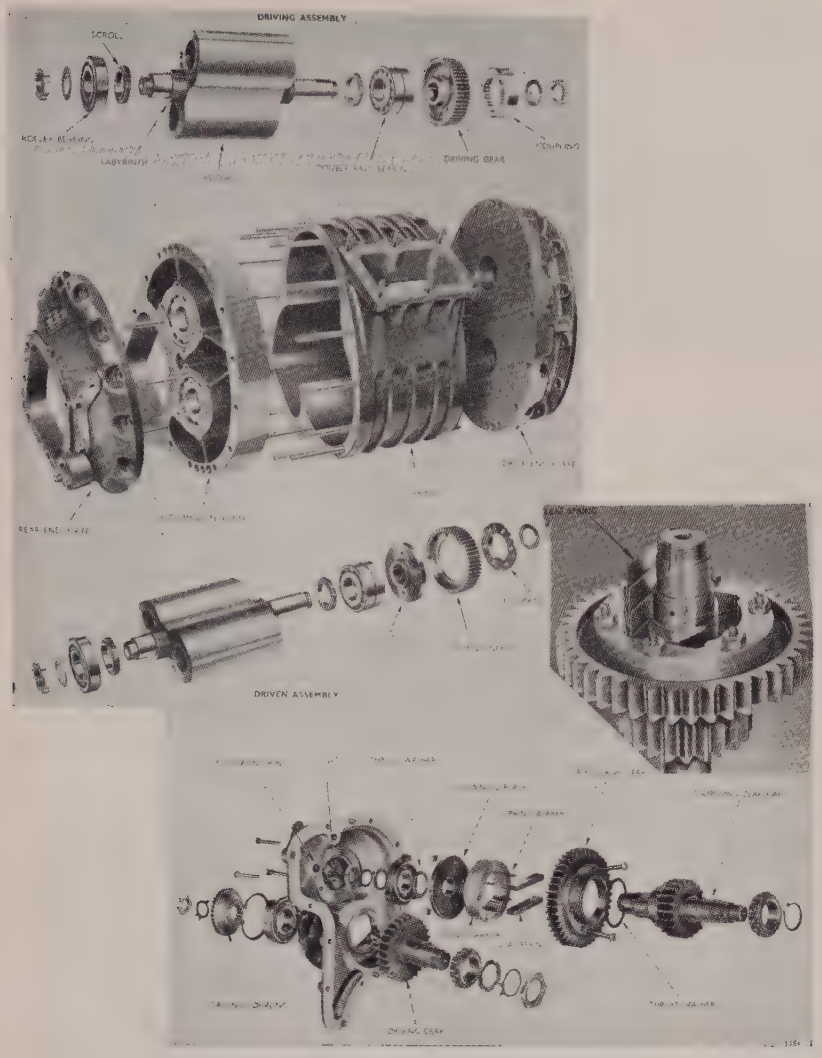


Fig. 11 — Rootes supercharger and drive

When a supercharger is fitted, this is driven through a torsionally flexible drive to dampen the torsional vibrations in the gear train. (See Fig. 11.)

To ensure concentricity of the crankshaft damper assembly, the hub of the damper is located on two cones (one bronze and the other steel) and driven by splines. (See Fig. 12.)

Cooling System

Due to high thermal efficiency, these diesel engines have a relatively low rate of heat rejection to the coolant.

The coolant is circulated through the oil heat exchanger, the top of the crankcase and the cylinder heads by a centrifugal pump. The pump casing also contains the thermostat and bypass which controls the rate of coolant flow through the radiator and assists rapid warmup. Passing into the cylinder block via the full length gallery, the coolant flows around the cylinder liners and then passes into the cylinder heads through transfer bobbins with O-ring seals. The main stream of the coolant flow is directed to the upper and hotter parts of the liners and a baffle is cast into the crankcase to ensure a high velocity flow around the top of the sleeve. As the water is introduced tangentially into this annulus, a high degree of scouring takes place.

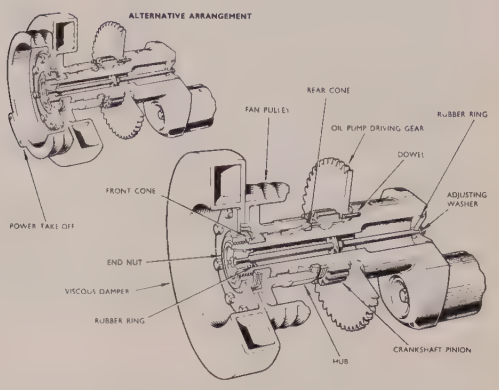


Fig. 12 — Viscous crankshaft damper mounting arrangements

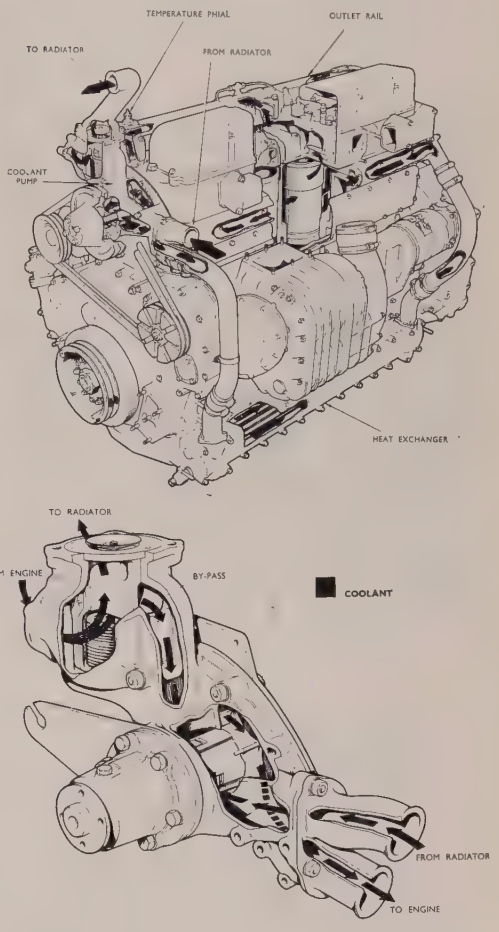


Fig. 13 — Coolant circuit

Below the baffle the lower part of the liner is cooled mainly by thermal syphon effect. After passing through the cylinder heads, coolant is fed back to the thermostat housing by an external rail. (See Fig. 13.)

Cold Starting

The direct-injection combustion system has a low area/volume ratio which directly assists cold starting. This feature, coupled with controlled air movement, high fuel atomization, and small valve overlap, has resulted in exceptionally good cold-starting properties. Additionally, an extremely

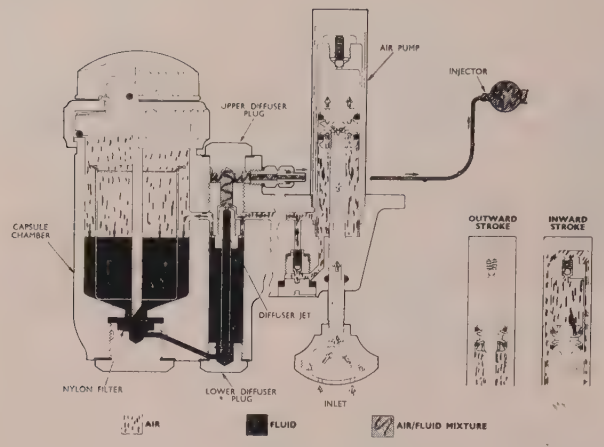


Fig. 14 — Start-pilot ether carburetor

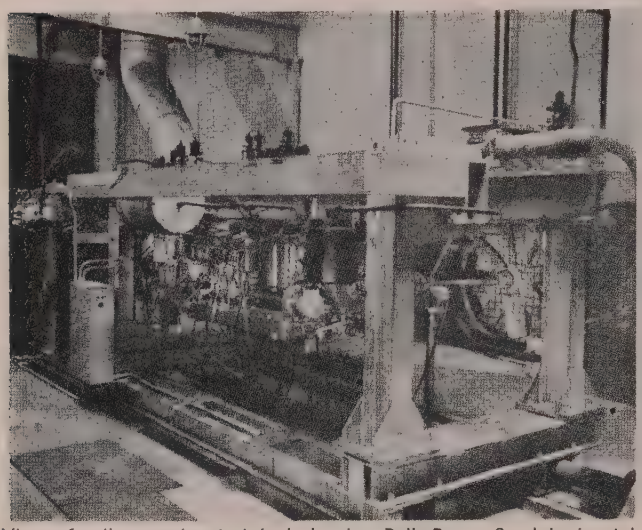
high torque starter motor is used which ensures adequate cranking speed even at sub zero temperatures.

The exceptional cold-starting ability of these engines can best be demonstrated by saying that they can start without any aids down to temperatures as low as 5 F and, with a simple ether device, start down to the limiting temperature of lead/acid batteries. Beyond this point additional aids can be provided to ensure satisfactory starting at – 65 deg after soaking at this temperature for two or three days. (See Fig. 14.)

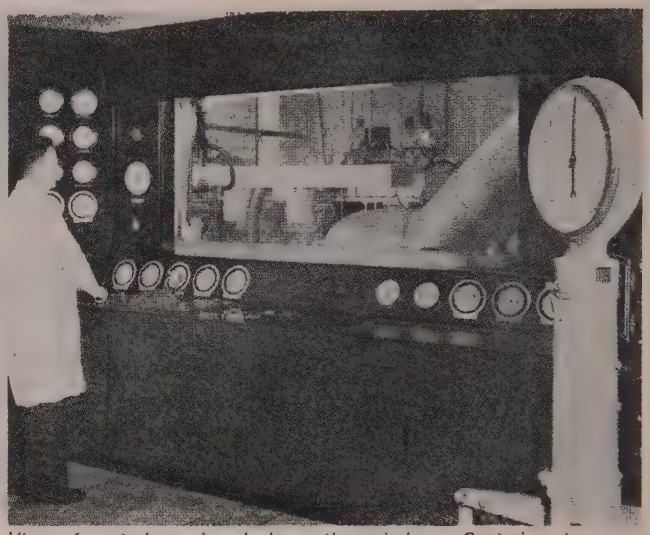
Environmental Testing

In closing I would like to mention a few phases of our development testing and, in particular, environmental testing.

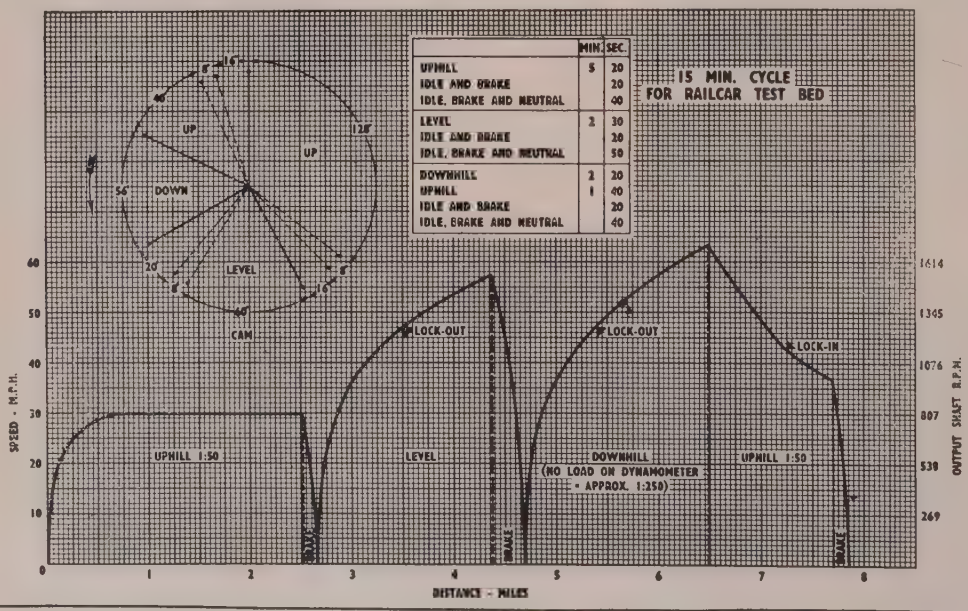
This type of testing plays an important part in our engine development program and a great deal can be done to reproduce some service conditions in



View of railcar engine test bed showing Rolls-Royce 8-cyl horizontal diesel and torque converter installation



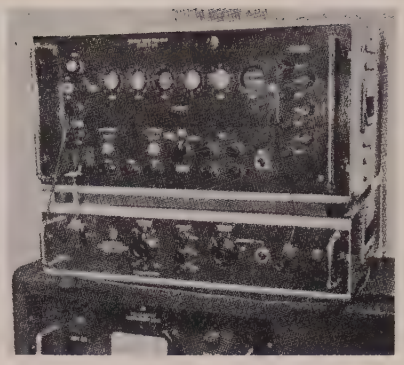
View of control panel and observation window. Control system comprises air and electrical circuits



A 15 min. test cycle is employed for cycle testing of the railcar power units. A motor drives a control cam revolving at four times per hour. The purpose of the cam is the operation of the dynamic brake at differing loads corresponding to 'up', 'level', 'down' and 'brake' on an actual rail section and the selection of idle and neutral positions between each stage. The cam carries a series of buttons in positions to give the correct sequence of timing for a 15 min. cycle. The engine can be stopped manually in any position in the cycle, a reset device on the cam allowing it to be indexed back to the initial position for starting

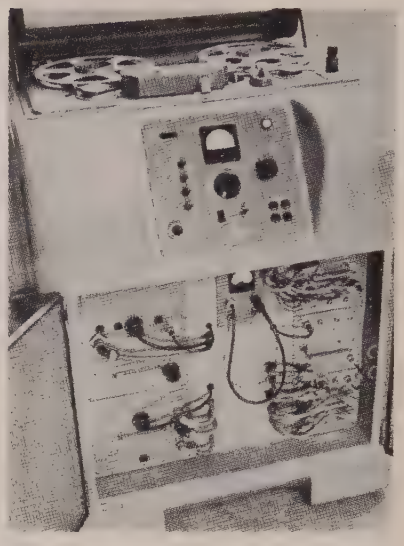
Fig. 15 — Railcar engine test rig

Scenes during cold-room trials of Rolls-Royce generating sets, simulating arctic conditions



Speed-indication equipment

General view of electronic panel, with more close-up view on right showing oscilloscope



Tape recorder set used in laboratory

Fig. 16 — Central electronics laboratory

the test cell. Testing, for example, may be carried out over a wide range of ambient air temperatures from −65 to +140 F or we may choose to introduce controlled amounts of airborne abrasives into engines. As a further step, a full-scale test rig may be constructed and I would cite our railcar engine test rig as an example of what can be done in this direction. (See Fig. 15.)

This test rig consists of a section of the railcar frame with the engine installation complete with cooling system and torque converter transmission, which is in turn coupled to a dynamometer. Three heavy flywheels are driven in series with the dynamometer to produce an inertia equivalent to half that of the railcar. Half the inertia value is chosen because in this particular engine installation, two engines are normally used. A brake with a high capacity is used so that the energy of the flywheels can be destroyed quickly; at one point in the operating cycle this energy amounts to 875 bhp. The system of controls enables the conditions of railcar service to be simulated on the test bed, actual railcar controls being used as far as possible to allow the maximum testing of equipment. The system is arranged for either manual or fully automatic control, in the latter case typical railcar route runs can be cycled for as long as required, reproducing the conditions of starting, accelerating, cruising—both up and down grade, decelerating, stopping, and idling.

I would also like to mention our method of utilizing electronic equipment in our test program. (See Fig. 16.)

In this present day when more and more electronic equipment is being used for analyzing test results, it becomes quite a problem for the technician to do his job satisfactorily amid all the distractions of the test cell itself. With this in mind we have recently installed a system at our plant in England whereby we have a central electronics laboratory devoted to the analysis of engine and rig test results. Each test bay is equipped with a multiple outlet junction box connected by a trunk cable system to the laboratory. The laboratory is in direct connection with the test cell by an intercom system. The provision of similar junction boxes in the laboratory enables a multiplicity of electronic equipment to be plugged into the test cell according to the requirements of the test in process. In this way the technicians are able to analyze the results in isolation and do not have restrictions upon the amount of equipment which may be used for any individual tests. An important adjunct of the specialized testing which is carried out in this laboratory is the use of tape recorders to make permanent records of the more problematical results. Records can be played back into the laboratory panel when required and this is particularly advantageous when it is necessary to re-examine the test results at a later date following new developments. The tape recorder in effect enables the electronic signals to be stored for repeat performances. By virtue of the transformation of test readings into a varying voltage, they can be continuously recorded by the instrument. When the record is played back to the oscilloscope, it is so synchronized as to display the test variations (that is, pressure, moment, strain, and the like) on a crank angle basis.

Figs. 17 and 18 will be found on p. 110.

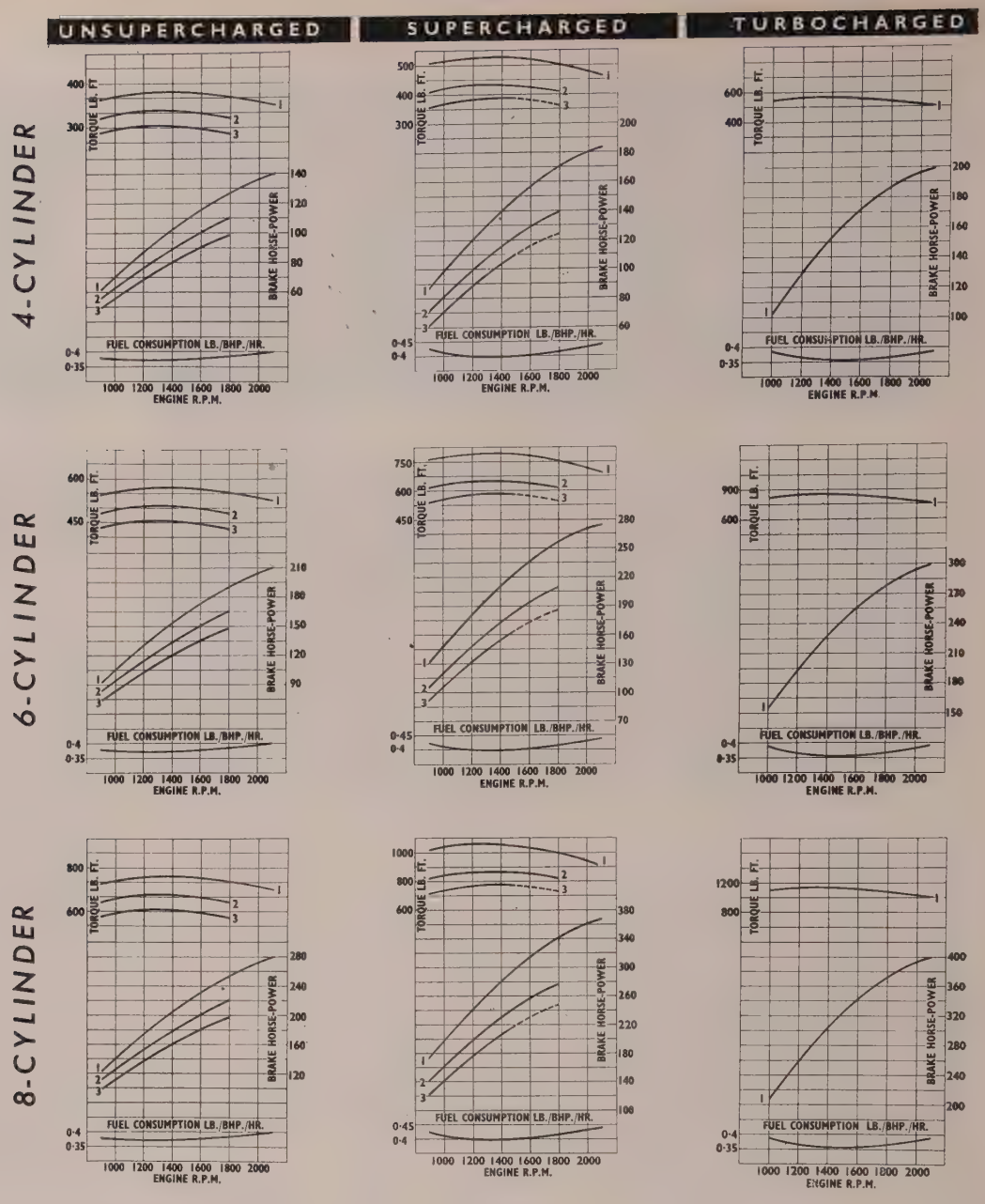
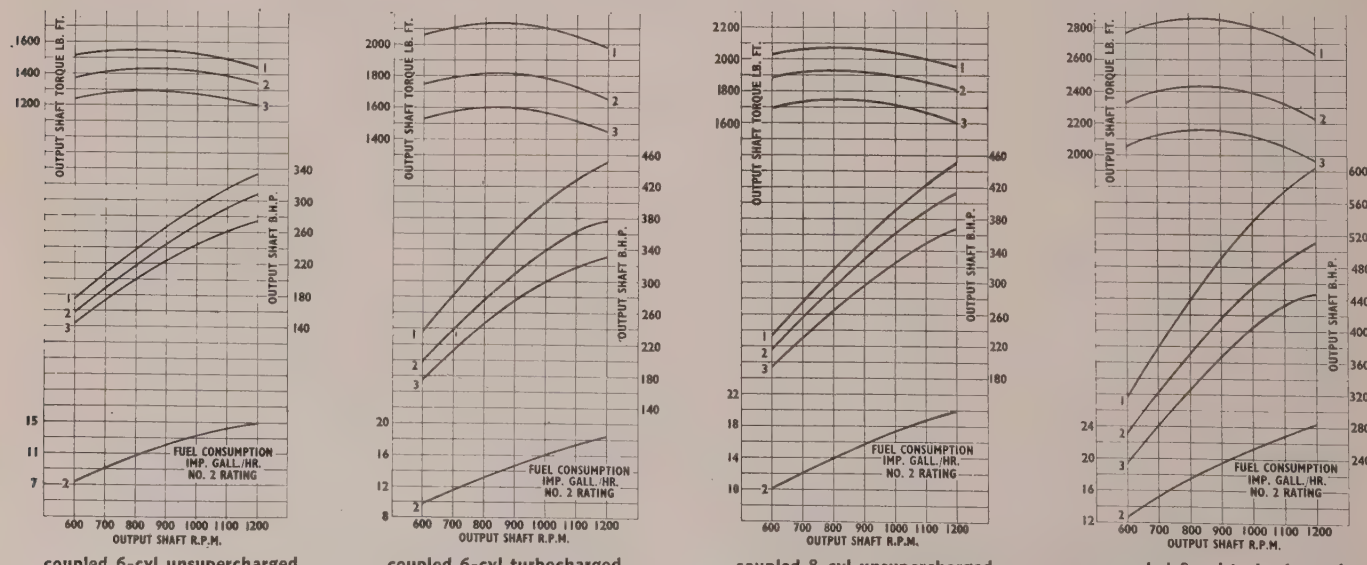


Fig. 17 — Single engine unit performance curves



coupled 6-cyl unsupercharged | coupled 6-cyl turbocharged | coupled 8-cyl unsupercharged | coupled 8-cyl turbocharged

Fig. 18 — Coupled engine performance curves for coupled packs with normal gearbox ratio 1.65/1

Non-Air-Breathing Auxiliary Powerplants

D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, Aerojet-General Nucleonics and A. M. Taylor, Aerojet-General Corp.

THE GENERAL problem of non-air-breathing auxiliary power systems suitable for missiles and space vehicles is considered. On the basis of minimum weight it is shown that for powers less than a few kilowatts, systems such as batteries, fuel-cells, solar cells, and radioisotopes are preferred. The particular choice depends mainly upon the required duration.

For powers larger than a few kilowatts, either chemical propellant-mechanical conversion systems or nuclear power systems must be used. Due to shielding considerations the nuclear auxiliary power systems are primarily restricted to satellites and space vehicles. The chemical propellant systems are required for durations of a few hours and less; the nuclear systems must be used for times greater than a few hours. Various types of chemical propellant-mechanical conversion systems are described.

The problems involved in a nuclear power system are considered in detail and the characteristics of a typical advanced state-of-the-art system are described. This system operates on the Rankine-liquid-vapor condensing cycle. Thermodynamic properties are given for several fluids that might be considered for such a system.*

IN RECENT years there has been an increasingly large amount of interest in non-air-breathing power systems. Missiles of all types from the relatively small air-to-air systems to the very large ICBM's require auxiliary power for the operation of the controls, guidance mechanisms, data telemetering systems, and the like. The satellites which we have sent aloft have required power for instrumentation, control, and communications. The satellites of the future will require power not only for instrumentation and communications, but also for environmental and comfort control for man and for propelling the vehicle in space. In fact, the "auxiliary" powerplant for satellite and space vehicles becomes the main powerplant when the vehicle is in its space environment.

The range of power versus duration combinations for such non-air-breathing powerplants is very large and the type of power required is also quite varied, running from simple mechanical translation, to rotating shafts, to pressurized hydraulic fluids, and finally to electrical power of various frequencies and voltage. There are systems requiring only fractions of a horsepower or kilowatts for periods of only a few seconds on up to satellite and space-type vehicles for the future which will require powers on the order of magawatts for durations of a year or more. The type of power system best suited to supply these needs can vary from batteries to fuel cells, solar cells, chemical-mechanical conversion systems, radioactive isotope sources, and nuclear reactor power systems.

The selection of a given system depends upon the particular vehicle, the already available energy

* Paper presented at SAE National Aeronautic Meeting, New York, April 1, 1959.

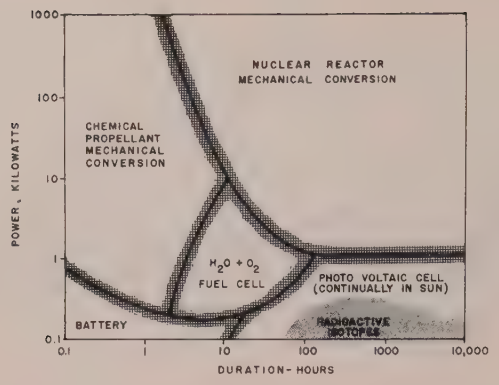


Fig. 1 — Power duration spectrum for various types of power systems available for use in space vehicles

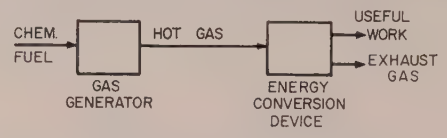


Fig. 2A — Schematic of chemical propellant-mechanical conversion power system

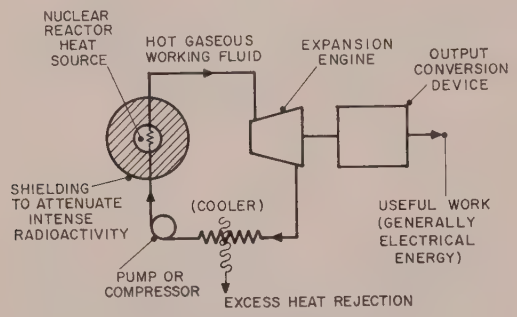


Fig. 2B — Schematic of nuclear reactor-mechanical conversion power system

sources, the environment, and the mission to be performed. There are, however, some basic desired characteristics common to all of the systems. In every case, one requires a minimum weight, minimum volume, maximum reliability, and satisfactory compatibility with other components of the vehicle system. Generally, take-off weight is many times the payload weight, and in the case of satellites, this ratio may be several hundred and more. Moreover, for satellites it presently costs about $500 for every pound placed in orbit. Thus, the weight of the auxiliary powerplant is worthy of detailed consideration in the overall system design. However, any reductions in weight must be commensurate with the required reliability of the system.

Areas of Application of Various Types

A qualitative and semi-quantitative comparison of the area of applicability of various types of power systems, in terms of minimum weight for a given power and duration, is illustrated in Fig. 1. Only mechanical conversion systems have been considered for the chemical propellant and nuclear energy sources. Direct conversion of thermal (heat) energy is very attractive in concept, since only a few or perhaps no moving parts would be required. This should result in an inherent power system simplicity and reliability. Considerable effort is being directed to systems of this type and, in fact, there are workable systems in existence. However, at present, these systems are all relegated to very low power outputs, low efficiencies, and low temperatures. Before they become attractive for large power production, some major breakthrough will be required. Eventually, direct conversion systems may be more desirable than conventional mechanical conversion systems, but this will not occur in the immediately foreseeable future.

The important conclusions to be drawn from Fig. 1 is that the battery, fuel cells, solar cells, and isotope energy sources are limited to systems of less than a few kilowatts of power. For powers greater than a few kilowatts, either chemically fueled or nuclear reactor mechanical conversion systems must be used. The particular type of plant depends upon the duration of the intended mission. The latter two types are considered in the present paper and it can be noted that they cover the major portion of the power duration spectrum. Some of the basic characteristics distinguishing chemical propellant-mechanical conversion systems from the nuclear mechanical conversion systems are discussed below.

Comparison of Chemical and Nuclear Systems

Chemical propellant-mechanical conversion systems obtain energy from the chemical reaction of one or more chemical propellants. The amount of energy available from a chemical reaction is many orders of magnitude less than that available from the nuclear fission process. The energy released per molecule in a chemical process is about one-half an electron-volt per molecule whereas about 200 mev per molecule are released in a fission process. Thus, for very large powers and long durations the amount of fuel required for a chemical system becomes prohibitive just from a weight and volume point of view. On the other hand, for a nuclear system the amount of fuel actually burned even for large powers and durations is so small that in essence only the weight of a critical mass is required for any power level. (This depends somewhat on the type of reactor employed.)

A schematic block diagram for a chemical system is shown in Fig. 2A. Basically, the chemical propellant-mechanical conversion systems involve a hot-gas generator and an energy conversion device. When the energy has been extracted from the gas, the gas is removed from the system. Thus, the fuel is expended and the overall system weight is reduced. This has important implications in connection with landing weights for the vehicle system.

A schematic nuclear power system is illustrated in Fig. 2B. It is important to note that the nuclear energy source (the reactor) is used only as a source of heat. Thus, the heat from the reactor is used to add energy to the cycle working fluid. The fluid is then expanded in an expansion engine where work is extracted from it. It is then taken to a cooler where the cycle excess heat is rejected, and thence to the pump or compressor where it is brought back up to the maximum pressure and introduced once more into the reactor.

There are several basic differences between the nuclear system and the chemical system described above. First, there is no fluid or material rejection from the nuclear power system. Thus, powerplant

weight remains constant, resulting in heavy landing loads for the vehicle (assuming that it is desired to land with the power system). Second, a means must be provided for cooling the working fluid; that is, extracting the cycle waste heat. This presents a very serious problem in the application of nuclear energy to space because all of the excess heat must be rejected by thermal radiation (Stefan-Boltzmann T^4 Thermal Radiation). This is unfortunate because in order to maintain reasonably sized radiators, very high sink temperatures are required. Another extremely important difference between nuclear and chemical systems not illustrated in Fig. 2B is that the nuclear reactor must have a certain critical mass in order to develop any power whatsoever. For this reason the reactor heat source is undesirable for very low powers (below 1–2 kw). On the other hand, it becomes a major advantage for higher powers since once a critical mass is obtained, it can be operated to very high power densities, the only limitation being that of heat transfer and materials. Still another factor of extreme importance in nuclear systems is the intense radioactivity associated with the nuclear fission process that occurs in the reactor. Even if no man is present on the vehicle system, some shielding or attenuation of these radiations is required since they can damage communication systems, materials, and the like, to the point where they are not operative. When man is aboard the system, the shielding required to attenuate these radiations to acceptable levels becomes extremely large and heavy. Thus, shielding weight limits the area of applicability of the nuclear power system. This is discussed in more detail in later sections.

Chemical Propellant Auxiliary Powerplants

Requirements

It has already been pointed out that the most significant requirements for APS's (Auxiliary Power System) are reliability, weight, storability, and compatibility with the other components of the vehicle system. In very short-duration missile applications, fuel weight is so low that APS efficiency becomes a secondary factor. In the APS's for the medium-duration earth-bound missiles (ICBM), the weights of the hardware and the fuel are both significant, and a proper balance between these two factors must be attained. In chemical APS's used for "long-duration" applications, fuel weight predominates and thus the efficiency of the APS is more important than the size and weight of the fixed hardware.

Typical Chemical APS Systems

A typical chemical APS system consists of three fairly distinct subsystems, each of which performs a definite function. The function of the first of these subsystems is to provide hot gases at relatively high pressure. The function of the second subsystem is to convert available energy in the hot gases to mechanical energy. The function of the third subsystem is to convert mechanical energy into hydraulic, pneumatic, or electrical output power.

Gas Generation — Most chemical APS's are driven by gases produced by the combustion or decomposition of solid propellants, liquid bi-propellants, or monopropellants. Fig. 3 shows typical successful arrangements utilizing these types of propellants.

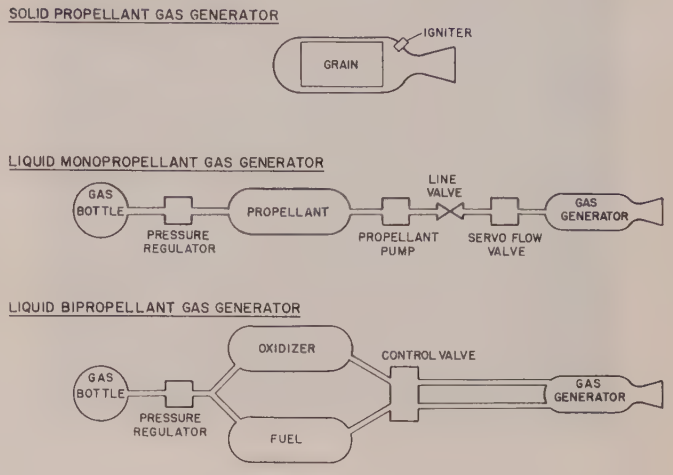


Fig. 3 — Typical gas generator subsystems

Solid Propellant — In solid propellant gas generators (Fig. 3) the feed sytsem and combustion chamber are combined. Such a gas generator has the inherent advantages of reliability, simplicity, and excellent storage capability. However, APS systems operating at partial load cannot use all of the gas generated by a solid propellant charge because the rate of gas generation is a function of chamber pressure, and of propellant temperature and composition, and not of the power demand of the vehicle system. Hence, at partial load, a high precentage of the gas must be ducted overboard unused.

Bi-Propellant — Fig. 3 also shows a bi-propellant gas generation subsystem. In this system, gas from a small pressurized tank is fed through a pressure regulator to force oxidizer and fuel from high-pressure tanks. These propellant elements pass through a liquid throttle valve into a chamber where combustion occurs. This type of gas generation system is normally used in place of the less complex monopropellant system only where sources of fuel and oxidizer are already available for main propulsion. Due to the relatively high complexity of the mixture ratio and ignition control components, the bi-propellant system has a lower inherent reliability than the solid propellant systems, as well as a poorer storage capability. However, in the bi-propellant systems, propellant flow can be metered in accordance with vehicle system power demands, with the result that specific fuel consumption at partial load is appreciably reduced.

Monopropellant — Fig. 3 also shows a typical gas generation subsystem using a monopropellant. This system is similar to the bi-propellant system except that it is simplified by the elimination of one fluid, and thus the need for mixture-ratio and ignition timing control. The system shown includes a propellant pump which is driven either by an electrical motor or by a shaft from the gearbox, but this pump can be replaced by a pressurized feed system similar to that shown for the bi-propellant system discussed above.

The monopropellant system, like the bi-propellant system, has the advantage that its propellant flow rate can be metered to meet variations in the power demands. Although it is somewhat more complex than a solid propellant gas generation system, it can

Fig. 4 — Typical APS turbine wheel

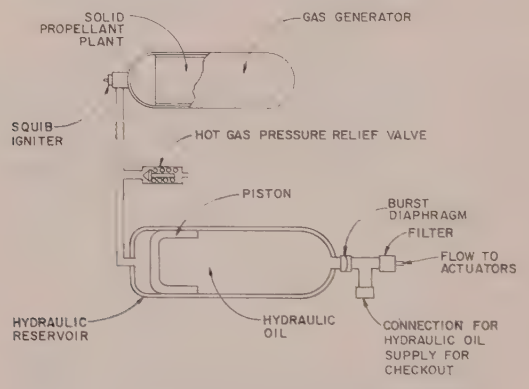


Fig. 5 — Schematic of hot-gas pressurized reservoir system

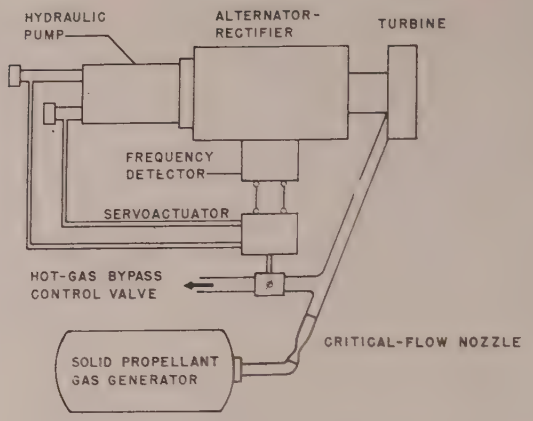


Fig. 6 — Schematic of solid propellant APS (hot-gas bypass control)

Table 1 — Propellant Performance Comparison

Propellant	Flame Temperature, T_c-F	Characteristic Velocity, C*-fps	Specific Fuel Consumption: Perfect Energy Conversion, lb/shp-hr	Specific Fuel Consumption: Representative State of Values, lb/shp-hr
LOX/RP-1	1500	3100	3.06	10.1
Hydrogen Peroxide	1300	3100	3.06	10.1
Normal Propyl Nitrate	2100	3500	2.26	7.8
Ethylene Oxide	1800	3600	2.26	7.5
Unsymmetrical Dimethyl-Hydrazine	1300	3200	2.16	8.6
Hydrazine	1500	4200	1.95	5.8
Solid Propellent	1850	3600	2.44	7.7
	2150	3900	2.08	6.5

* 50% turbine efficiency, 100/1 pressure ratio.

be developed for each application to perform with a high degree of reliability.

Conversion to Mechanical Energy

Turbines — Most of the chemical APS's currently in use are driven by turbines. This is due to the relative simplicity of the turbine unit, which is particularly suited to long duration operation because of its small number of moving parts.

The conditions under which an APS turbine must operate differ significantly from those of a stationary steam installation or an air-breathing gas turbine. The APS turbine must produce low power from high-temperature gases at high pressure ratios. To operate at a low power level, the turbine must have a small arc of admission, with the result that there is a reduction in efficiency due to partial admission losses. In addition, the ratio of blade velocity to gas-spouting velocity is considerably lower than that required for peak efficiency, even with turbine wheel tip speed of 1600–1700 fps. The result is that most of today's single-stage APS turbines (Fig. 4) perform at efficiencies of 30–40%.

Hot-Gas Motors — The hot-gas motor is another machine now under development for the conversion of hot-gas energy to mechanical energy. Because of the particular suitability of positive displacement machines for operation at low specific speeds and low power, they are expected to be competitive with the turbine in the low power range. Several types of positive displacement motors are under development for use in chemical APS applications, including reciprocating piston motors, vane motors, and other types of rotating motors. Consideration is being given to designs in which pressurized hot gas is admitted throughout the complete power stroke, and to designs in which the hot-gas feed is cut off before completion of the power stroke in order to utilize some of the internal energy of the gas by expansion. This latter type of motor is expected to be more suitable for use for longer duration applications where efficiency is of greater significance. A disadvantage of hot-gas motors is that they contain a relatively large number of surfaces which rub at high temperatures, causing excessive wear.

Hot-Gas Pressurized Reservoir — The hot-gas pressurized reservoir system (Fig. 5) is a means of providing constant pressure hydraulic power directly from the hot gases supplied by the combustion chamber. This system is used for short-duration applications where the oil can be discharged overboard as it is used, and where no electrical power is required. In addition to a hot-gas generation subsystem, this system includes a hot-gas pressure regulating valve, and the reservoir which contains a free piston and the hydraulic oil supply. During operation the hot-gas displaces the hydraulic oil at constant pressure.

Hot-Gas Actuation — The hot-gas actuation system converts the available energy of hot gas directly to mechanical energy. This system is used primarily

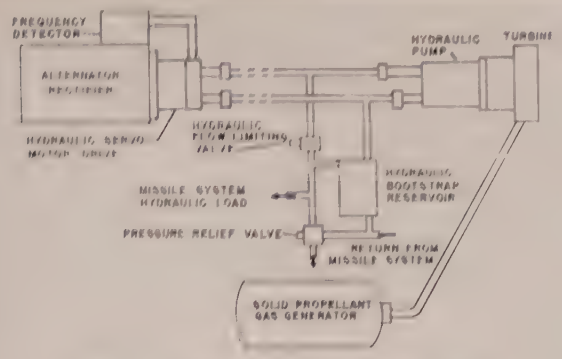


Fig. 7 — Schematic of solid propellant APS (flow-limiting hydraulic flow control)

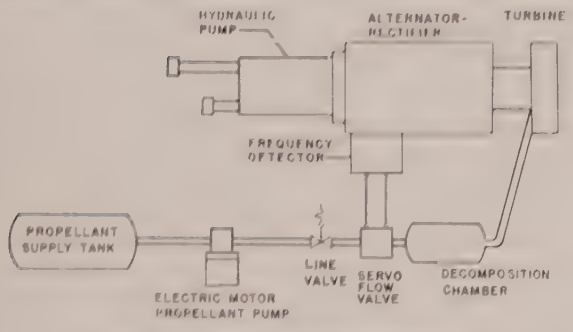


Fig. 8 — Schematic of monopropellant APS (liquid throttling control)

for the actuation of missile control. In this system hot gas, supplied by the gas generation subsystem and held at constant pressure by a regulating valve, passes through a servo valve to a double-acting actuator. Hot-gas actuation systems are inherently simple in that they do not require hydraulic fluid. The primary disadvantage of this system is development of its relatively complex hot-gas components.

Conversion to Output Power — Hydraulic pumps, alternators, and occasionally, pneumatic blowers, are used to convert mechanical energy developed in the APS to the type of power required by the vehicle. Normally, 3000-psi hydraulic power and 400-cps, 3-phase, 115-volt a-c power with voltage and frequency regulation of 1.5% are required. It is anticipated that future missile APS systems using miniaturized components will be designed for higher pressure hydraulic power and higher frequency (1600–3200 cps) electrical power.

Typical Chemical APS Systems

The subsystems previously discussed can be combined into a variety of complete APS systems, each of which is suited to the particular requirements of its application. Since it is not feasible to present all of the possible combinations in this paper, several types which have been successful have been chosen for further discussion.

Solid Propellant Hot-Gas Bypass APS System — Fig. 6 is a schematic diagram of a solid propellant APS system which produces electrical and hydraulic power. The speed of this unit is controlled to provide 400-cps frequency by a control system which bypasses excess gas as the output load is varied. The gas generator chamber pressure is maintained relatively constant through the use of a critical flow nozzle upstream of the control valve and turbine.

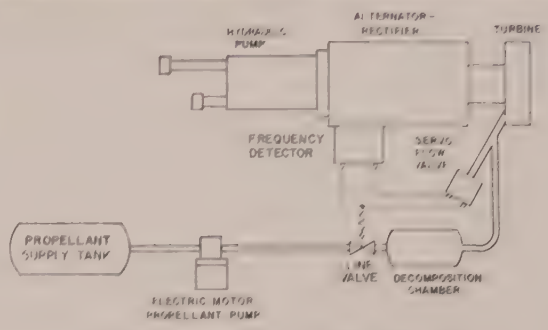


Fig. 9 — Schematic of monopropellant APS (gas throttling control)

The primary advantage of this system is the simplicity of the gas generator subsystem. The primary disadvantage is that the gas generator produces gas at the maximum power level at all times and dumps excess gas overboard under partial load conditions.

Solid Propellant Hydraulic Flow Control APS System — Fig. 7 shows another system which produces electrical and hydraulic power. In this system, all hot gases from the gas generator pass through the turbine, which in turn drives a fixed-displacement hydraulic pump. Some of the hydraulic flow powers a hydraulic-motor-driven alternator, with alternator frequency controlled by a hydraulic servo-valve. Since the alternator operates at a constant speed, the hydraulic flow to this portion of the system is constant. The remaining portion of the hydraulic fluid passes through a flow-limiting valve to the missile hydraulic load system and back to pump suction. The function of the hydraulic flow-limiting valve is to control the turbine speed when variations in propellant bulk temperature cause the gas flow to vary. This system also has a pressure relief valve in parallel with the missile load to provide a constant pressure in the hydraulic fluid upstream of the load. A bootstrap type of hydraulic reservoir provides pump suction pressure which is directly proportional to the pressure at the hydraulic load. This system, like the hot-gas bypass system, must produce gases at the full-power level at all times, and thus wastes energy when the system is under partial load. However, the system is well-suited for use on relatively short-duration missile applications, where propellant weight is of secondary importance.

Monopropellant Liquid Throttling APS System — Fig. 8 is a schematic diagram of a liquid monopropellant APS system supplying electrical and hydraulic power. Hot gases from the decomposition chamber pass through the turbine which drives the alternator and hydraulic pump. Alternator frequency and turbine speed are controlled by a speed control, which operates a servo valve to regulate propellant flow. The advantage of this system is that it provides the correct amount of propellant to keep the alternator at constant speed regardless of the load demand. Its main disadvantages are the relative complexity of the gas generator and propellant feed system, and the partial loss of response to load changes, due to the storage of energy in the decomposition chamber.

Monopropellant Hot-Gas Throttling APS System — Fig. 9 is the schematic diagram of a liquid monopropellant system in which frequency control is obtained by throttling the hot gases at the turbine

nozzle. This system provides a faster response than the liquid throttle control system by permitting the decomposition chamber to operate at reasonably constant pressure levels.

Propellant Performance

Table 1 shows the typical performance of several monopropellants, bi-propellants, and solid propellants which are available for use in APS systems. The parameter, characteristic velocity (C^*), is a partial measurement of the available energy per pound of a propellant. It is a function of the generated gas properties-molecular weight, ratio of specific heats, and temperature. It is desirable to select a propellant which has a high value of C^*. The table shows that the C^* of the monopropellants ranges from 3100 fps for hydrogen peroxide to 4200 fps for hydrazine at 1500 F. The C^* values of the two solid propellants shown are approximately in the middle of this range, while the C^* for the liquid oxygen/RP-1 combination is near the bottom of the performance range.

Limits of Chemical APS Performance

Table 1 also lists two values of specific fuel consumption for each of the propellants considered. The first of these, entitled "Perfect Energy Conversion," is based on the assumption that the efficiency of the turbine and output components are 100%, and that the pressure ratio across the turbine is infinite. Although such minimum fuel consumptions cannot be reached in practice, they provide a yardstick for comparison of actual performance with ideal performance.

Actual APS Performance

The second specific fuel consumption listed in Table 1, entitled, "Representative State of the Art Values," is based on more realistic turbine performance and pressure ratios. In calculating these values, the pressure ratio across the turbine nozzle is taken as 100/1 and the turbine efficiency is assumed to be 50%, a value which can be realized by multistage or re-entry turbines, even at moderately low load conditions. Thus, specific fuel consumption of 5.8 lb/shp-hr is representative of values attainable with combinations of equipment now available. With advances in the metallurgy of high-temperature materials and with expected improvements in turbine design, this value can be reduced to approach the limits described above more closely.

Comparison with Air-Breathing Systems

The specific fuel consumption presented here for chemical APS's appears to be very high in comparison with those of existing air-breathing engines; however, it should be remembered that these units must carry along all of the fluids (fuels) necessary for operation of the unit. A more reasonable comparison can be made if fuel consumption is calculated for air-breathing engines based on the total weight flow of air and fuel. On this basis an aircraft piston engine with a specific fuel consumption of 0.45 lb per bhp-hr requires approximately 7.2 lb of air and fuel per bhp-hr. This value is more consistent with the performance figures shown for chemical APS's. In making this comparison, however, it should be pointed out that there are compensating differences between air-breathing and non-air-breathing systems which should be kept in mind. The air-breathing engine operates at lower pressure ratios and higher temperatures than the chemical APS. Moreover, the chemical APS must operate at a partial admission in most low power applications resulting in a "built-in" inefficiency.

Nuclear Auxiliary Powerplant

In this discussion the salient features of space nuclear powerplants are considered and their general characteristics described. Following this are more detailed discussions of the components of the system.

The schematic configuration of a nuclear power system is illustrated in Fig. 2B and has been discussed previously. Briefly, the system consists of the nuclear reactor (heat source) a working fluid, an expansion engine (to extract energy from the working fluid), an output conversion device (to transform the expansion engine work into the desired form), and a cooler (which removes the cycle excess heat and rejects it from the system). For space uses, the "cooler" is actually a radiator which dissipates the heat by thermal radiation.

Nuclear reactors can be classified in many ways, but one method universally used depends on the neutron energy and separates reactors into thermal, epithermal, or fast reactors. The names are derived from the energy level of the neutrons that are most predmoninant in maintaining the chain reaction in the reactor. In terms of electron volts, the range of neutron energy varies from the low or thermal level of 1/40 ev to the very high value for fast neutrons of 2–5 million electron volts (mev). Epithermal neutrons are those in the middle portion of this range. The physical characteristics of these three types of reactors may be quite different, although the fission process itself is basically the same. This results in large variations in reactor size, depending upon the reactor type, as well as on power level. For the present power range of interest (from 10 to 1000 kw electrical power output), fast neutron reactors have been found to be smaller than epithermal reactors. (Thermal reactors are larger still than epithermal reactors.) Because size and weight are particularly important in space systems, the fast reactor is the subject of much attention.

General Configuration and Geometry — The radiation of high-energy neutrons and gamma rays associated with the nuclear fission process presents a major problem in the design of nuclear auxiliary powerplants. A fast neutron reactor capable of producing 10-kw net is less than a foot in diameter and weighs about 400–500 lb. The shielding required to reduce the radiation from this reactor to a level acceptable to man is about 5–6 ft thick and weighs approximately 50,000 lb, if a spherical shield that completely encloses the reactor is used. Adding to this the necessary power-conversion equipment results in a powerplant that weighs between 50,000 and 60,000 lb and produces only 10-kw output. It is evident that the only way in which reasonable system weights may be obtained for low-power-level nuclear plants is to somehow reduce the shield weight. This is possible only if a partial or "shadow" shield can be used. Such an arrangement is illustrated in Fig. 10. In this arrangement, the crew and other radiation-sensitive system components are located in the shadow projected by a partial shield. Typi-

cally, only 3–5% of the total area around the reactor need be shielded, thus, reducing the shield weight as much as 95–97%. However, this method works only where there is nothing nearby that would scatter radiation, so that it would reach the crew by an indirect path. In other words, simple shadow shielding is limited to locations such as outer space where there is no atmosphere to scatter radiation. For applications of nuclear auxiliary powerplants to manned vehicles that are in the atmosphere, the shielding must completely surround the reactor.

Separation of Powerplant and Vehicle — In applying nuclear energy to powerplants for satellites and space vehicles, weight is of prime importance because it presently costs about $500 per pound placed in orbit. Thus, every effort must be made to produce the lightest designs possible. Even though the shadow shielding is used, the shield still constitutes a significant proportion of the power system weight. In fact, for the low power levels (10–20 kw), this shield can constitute as much as 70% of the power system weight. One method of reducing weight is to separate the power system from the vehicle as illustrated in Fig. 10. This is effective because the shield thickness is proportional to the logarithum of the inverse of the square of the separation distance. In fact, for a large enough separation, no shield would be required. However, such large separations are generally not practical because the allowable vehicle maneuver rates with such a large separation would be very limited. If the vehicle and power system were connected by a long flexible cable, a separate control system on the powerplant would be required to keep it oriented with respect to the vehicle, and to keep it out of the flight path of other vehicles. If the long separation distances were accomplished by a boom-type structure, then excessive boom weights would result to prevent excess deflection of the power system, or damage to the booms in maneuvering. Typically, the power system can be extended anywhere from 30 to 200 ft from the vehicle, depending, of course, upon the particular requirements of the vehicle system. The variation of shield weight with distance is further discussed in the section on Reactor and Shield.

Radiator and Its Location — In space, excess cycle heat must be dissipated by Stefan-Boltzmann T^4 radiation. Because of this, the cooler (or radiator) is of paramount importance in the overall system design. For example, a powerplant with a 10% thermodynamic efficiency, and a radiator temperature of 500 F will require approximately 21 ft²/kw. Thus, even at 10–20 kw power levels, about 200–400 ft² of area is required. Moreover, this is effective area which must actually "see" space. If this radiator were a cylinder 2 ft in diameter, a length of 18–36 ft would be required. If the radiator temperature were 1000 F, then only 4 ft²/kw would be required. In other words, a five-fold reduction in size is gained by increasing the temperature 500 F. It is apparent that high radiator temperatures are required to obtain minimum radiator size and weight. This is especially true for high power levels, where the radiator becomes proportionally a larger part of the total powerplant. High radiator temperatures, of course, require high source temperatures and, thus, make the reactor design more difficult.

In locating the radiator, the requirement that all system components be confined within the shadow projected by the shield must be considered. If a component protrudes beyond this shadow, it can scatter enough radiation back to the vehicle to be harmful to the crew. Even at the low power levels, the required radiator area (and necessary volume) may be too large to permit placing the radiator on the reactor side of the shield, or to use the exterior surface of the shield as a radiator surface. Then the skin of the vehicle must be used as a radiating surface, or the radiator must be placed between the shield and the vehicle. It is usually not possible to use the vehicle skin because it is required for other functions; thus in many cases the only location which can be used for the radiator in manned systems is between the shield and vehicle. To prevent protrusion of the radiator elements beyond the shadow cone, the shield-vehicle separation must be large enough to provide sufficient "shadow volume" to contain the radiator.

Fission Product Leakage and Fluid Activation — Locating the radiator either on the vehicle or between the shield and the vehicle poses a new and important problem in plant design. The problem is radioactivity in the working fluid of the power system. In the "one-loop" system the working fluid passes through the reactor, the turbine, into the radiator, and back to the reactor via a pump or compressor. (See Fig. 11.) If this working fluid should become radioactive, it would subject the vehicle to large doses of radiation. The fluid could become radioactive in three ways: by the fluid itself be-

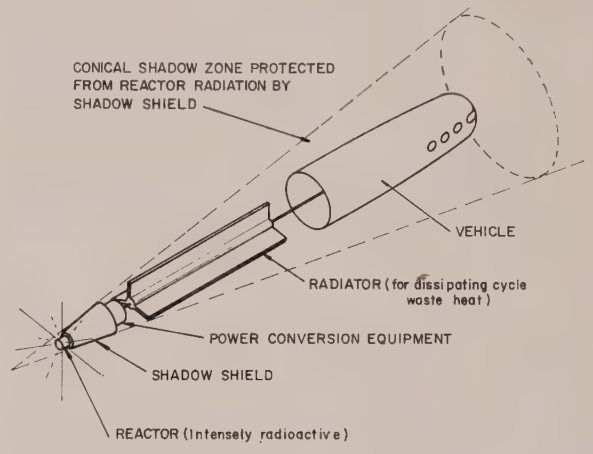

CONICAL SHADOW ZONE PROTECTED FROM REACTOR RADIATION BY SHADOW SHIELD
VEHICLE
RADIATOR (for dissipating cycle waste heat)
POWER CONVERSION EQUIPMENT
SHADOW SHIELD
REACTOR (Intensely radioactive)

Fig. 10 — General configuration of manned space nuclear power system and vehicle

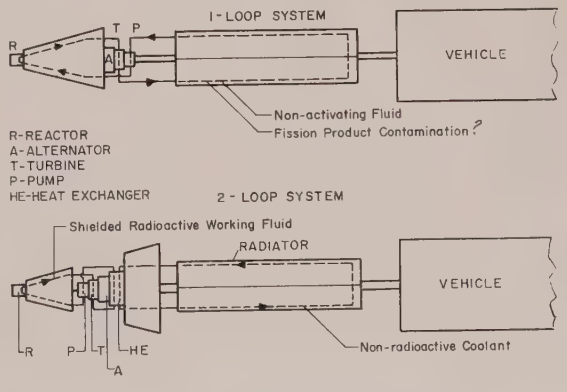


Fig. 11 — Fission product leakage and fluid activation

Table 2 — Typical System Characteristics

Maximum Fluid Temperature			1,600 F	
Sink Temperature			1,000 F	
Thermodynamic Efficiency			12%	
Radiator				
Shape Factor	0.8	Emissivity	0.8	
Area required per kilowatt electrical output — 5 ft²/kw(e)				
Shield				
Conical Shape				
	10 kw(e)		1,000 kw(e)	
	2-loop	**1-loop[a]**	**2-loop**	**1-loop**
Radiation Area	50 ft²	50 ft²	5,000 ft²	5,000 ft²
Separation Distance	30 ft	30 ft	150 ft	150 ft
Shield	1,300 lb	1,000 lb	3,000 lb	2,000 lb
Radiator	70	70	6,500	6,500
Reactor	500	500	1,500	1,500
Heat Exchanger	50	—	2,000	—
Rotating Machinery	90	80	3,500	3,000
Fluids	100	50	1,500	600
Miscellaneous (controls, structure, etc.)	150	150	500	500
Total	2,260 lb	1,850 lb	18,500 lb	14,100 lb

[a] Assuming the same cycle efficiency, radiator area, and the like.

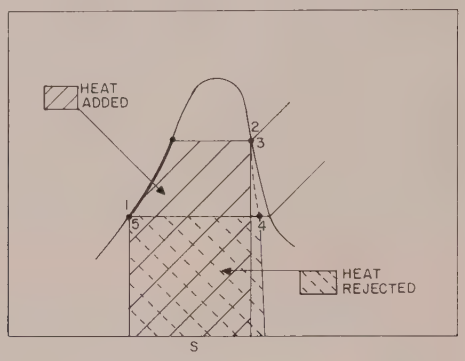


Fig. 12 — Simple classical Rankine cycle

1 - 2 HEAT ADDITION IN THE REACTOR
3 - 4 EXPANSION IN THE TURBINE
4 - 5 HEAT REJECTION & CONDENSATION IN THE RADIATOR OR (CONDENSER)
5 - 1 RE-PRESSURIZING IN THE CONDENSATE PUMP

coming activated as it is bombarded by neutrons in going through the reactor; by entraining the intensely radioactive fission products that might leak or diffuse into the fluid from the reactor fuel elements; and by neutron and gamma activation of particles of metal which have been corroded away from the fluid containment vessels and are carried in the fluid stream.

It is possible to select a working fluid that will not be significantly activated under neutron bombardment. Helium gas is such a fluid, and there are several others. The problem of guaranteeing that no significant amount of fission products becomes entrained in the fluid is an entirely different matter. Radioactive fission products are continuously formed when the reactor is in operation. If personnel were exposed to the radioactivity from the fission products formed in a few weeks operation of a low power reactor, they could receive a radiation dose rate as much as 10^6 times the rate that would be considered a reasonable maximum for a short duration exposure. This, of course, would be lethal. Thus, as much as 10^{-4}% of the fission products formed (in this example) cannot be allowed to escape into the working fluid of a one-loop system.

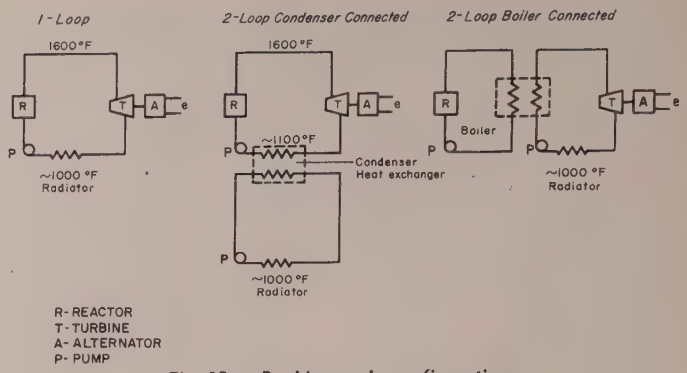


Fig. 13 — Rankine cycle configurations

This presents a rather formidable reactor design problem because a pinhole in a fuel element could easily allow this quantity of fission products to escape into the fluid. The high temperature required in these reactors further aggravates the problem. As fuel-element design and cladding-material technology advance it is reasonable to expect that the design requirement can be met. However, considerable development effort will be required.

The Two-Loop System — To guarantee that there will be no significant entrainment of fission products in the radiator fluid, the "two-loop" system can be used. Fig. 11 illustrates such a system. In this arrangement, the working fluid is confined to the smaller part of the power system, such as reactor, turbine, pump (or compressor), and heat exchanger (which transfers heat between the two loops). The shield is split into two parts and the power-conversion equipment and heat exchanger are located between them. If any significant entrainment occurred or even if a fuel element failed completely, there would be sufficient shielding between the working fluid and the vehicle to adequately protect the crew. The radiator fluid is isolated from the working fluid by the heat exchanger. A disadvantage of this system is an increased shield weight resulting from splitting the shield. This is so because even though shield thickness is about the same, the cross-section of that part of the shield closer to the vehicle is larger than for an unsplit shield. Further weight disadvantages of the two-loop system result from the circulating pump for the radiator and the additional heat exchanger.

The Rankine Cycle — It has been shown that excess heat from the power cycle must be dissipated by thermal radiation and thus high radiator temperatures are required to obtain radiators of reasonable size and weight. High radiator temperatures in turn require high source temperatures. Thermodynamic efficiency *per se* is not important, but thermodynamic efficiency from the point of view of minimum system weight is important. Many thermodynamic cycles can be postulated for use in a nuclear power system. After rather extensive studies the Rankine liquid-vapor condensing cycle has been found to be the most attractive for the conditions pertaining to space power systems, particularly for the low power level range. The simple classical Rankine cycle is illustrated on the *T-S* diagram in Fig. 12. The more important of the many reasons for this choice are: (1) Nearly all the heat is added and rejected at constant temperature; (2)

the back-work ratio (pump work) is relatively small compared to the turbine work, so that reasonable cycle efficiency is obtained even with relatively poor component efficiencies; and (3) the cycle is basically very simple, that is, only a few components are involved. More detailed comparisons of the Rankine liquid-vapor cycle and the Brayton gas cycle are given in the next section.

Zero Gravity Problems — The selection of a Rankine cycle is necessary from the powerplant weight, size and other considerations; however, the use of such boiling and condensing systems considerably complicates the design problem since the system must operate in a zero gravity environment. A major problem in this respect is the condensing of the vapor in the condenser or radiator. The boiling in the reactor (or boiler) can be expected to be a much less severe problem than condensing because high heat fluxes are possible and also high fluid density and velocities can be used. Thus, by use of high pressure drop, swirling vanes, and high velocities, the solution of the boiling problem should be relatively straightforward.

Several of the possible configurations of the Rankine cycle are illustrated in Fig. 13. The one-loop system is seen to be the simplest and involves boiling in the reactor and condensing in the radiator. The two-loop system may be either condenser-connected or boiler-connected. For the same maximum and minimum temperatures it is seen that the two-loop systems suffer a thermodynamic penalty over the one-loop systems by virtue of the necessary temperature differences in the coupling heat exchanger. If the radiator must serve as a condenser (such as in the one-loop or the boiler-connected two-loop system), a design conflict may arise. Radiation to space requires that the radiator area completely "see" space (a flat sheet is optimum in this respect). Condensation, however, may require production of local acceleration to control the vapor and liquid phases. This may be accomplished, for example, by spiral tubes, but the use of spiral tubes can significantly reduce the radiation space "seeing" factor. The condensing and radiation problems can be separated by the use of the condenser-connected system. In this system the coolant in radiator loop can be a single-phase fluid, thus having no condensing problem. The condenser, thus, becomes the coupling heat exchanger which can be a compact high heat flux component amenable to a geometry which is optimum for zero gravity condensation. In the condenser-connected system, the use of a liquid rather than a gas in the coolant loop would result in the greatest heat-transfer coefficients and the smallest pump work. This would, however, increase the weight of the fluid inventory. It is probable that the condenser-connected system offers fewer zero-gravity problems. However, the selection of a given system will depend upon the particular circumstances and conditions involved in the overall system in question.

Typical Nuclear Mechanical Conversion Power System Characteristics — Detailed specifications for a nuclear auxiliary power system, that is, operating temperatures, working fluid, component weights, and so forth, can only be determined after extensive investigations of the specific requirements such as power output and duration. However, for purposes of illustration, some typical weight characteristics are summarized in Table 2. These are based on a source temperature of 1600 F, sink temperature of 1000 F, thermodynamic efficiency of 12%, radiator area per kilowatt output of 5 ft^2, and a conically shaped shield. The shield weights given are of the conical shield, with the reactor separated from the dose point as indicated. The radiation dose at this point is equivalent to about 10 roentgen per month, which is an arbitrary value, but appears to be in the proper range. It is apparent at both power levels that the one-loop system offers definite weight advantages.

At the low power level, the shield constitutes more than 50% of total system weight, whereas the radiator weight amounts to only 5–8%. On the other hand, at the high power level, the shield amounts to only about 15% of the total; whereas, the radiator constitutes approximately 30% of the weight. In the low-power-level case, radiator volume is a controlling factor in system design. At the higher power levels, both radiator volume and weight are important factors. The weight advantage of the one-loop system over the two-loop system stems from several factors: elimination of the heat exchanger, reduction of the shield weight (since it does not have to be split), reduction of the rotating machinery weight (since fewer components are involved), and reduction of the total fluid holdup weight. Fluid is saved because only a fraction of the radiator volume is actually filled with condensed liquid, perhaps something like 30% of the total volume.

A sharp reduction in specific weight occurs with increases in power level. For the one-loop design, at 10 kw, the weight is about 225 lb/kw(e) whereas at 1000 kw the specific weight is down to about 14 lb/kw(e). Further large reductions in specific weight with increased power will not be obtained, because the shield no longer constitutes a significant portion of the total system weight.

These figures are only approximations, but do show the trends and typical weights to be expected in these systems. Refinements of design and more complete evaluations could very likely result in lower weights.

Another important factor to consider is the relationship of the vehicle configuration to the power system weight. A conical shield has been assumed in the present illustrations. However, it is shown in the sections on the reactor and shield that the configuration of the vehicle could possibly increase the shield weight by as much as 200% or more over the simple conical configuration. Obviously, this has a very important influence on the weights attainable in the power systems.

Unmanned Spaced Vehicles — In the previous sections on nuclear auxiliary power systems the stress has been laid on adequate shielding for protecting a man against the harmful radiations emitted by the nuclear reactor and associated system. However, in certain cases it may be desirable to have large amounts of power in unmanned space vehicles in which the only shielding necessary is to protect those parts of the vehicle instrumentation which would suffer radiation damage upon exposure to the gamma rays and fast and slow neutrons. Considerable investigation has been made on the effects of radiations upon various electronic and structural compounds. For example, see the Proceedings of

the Third Annual Radiations Effect Symposium, October, 1958. The situation may be summarized by the fact that materials vary greatly in the amount of radiation they can stand before deteriorating to where performance is not satisfactory. Among the most sensitive are found to be transistors, semi-conductors and organics, such as are used for insulation on wires, condensors, and other electronic components. In certain cases, doses as small as 10^6 roentgen or 10^{10} fast neutrons per square centimeter may effect performance of components so that systems such as a radio would fail. In order to avoid this extreme sensitivity to radiation (which, incidentally, is 10^4 times greater than human beings can stand) efforts may be made to "harden" the exposed equipment. This may be done by replacing semi-conductors by "hard" vacuum tubes which are some 10^4 to 10^5 times less sensitive to radiation. However, as a practical matter even with extreme care in the choice of materials and in the design of the equipment, it cannot be expected that a system can be built which will not suffer ill effects from radiation in the amounts of 10^7 to 10^8 roentgen.

If a system such as shown in Fig. 10 is used where the objects to be shielded from the eminating radiation are the sensitive electronic and other pay loads (rather than a man), calculations show that the shield weight may be reduced between a factor of 1/2–2/3, depending upon the exact composition of the pay load and the other parts of the vehicle. Thus, from Table 2 we can easily compute that the overall weight of a 10-kw plant would vary between 1250 and 1600 lb and for a 1-megawatt plant between 13,100 and 13,500 lb. It is interesting to note that for large powers, such as one megawatt, it makes little difference in overall weight of the system, whether one is shielding for man or equipment, since the shield weight is relatively small for such large powerplants.

Another interesting thing that should be noted about unmanned vehicle utilization of nuclear power is the possibility of shielding the equipment rather than the reactor. It is obvious that this would turn out to yield a lower total weight if the equipment package which must be shielded is smaller in dimensions that the reactor-system which would otherwise be shielded. Still another possibility exists in the rearrangement of the components from that used in the manned vehicle, since many parts of this structure (radiator, and the like) do not require any shielding, if they are fabricated of the proper materials.

Further Detailed Considerations of Nuclear Power Systems

In the previous discussions, some of the major features of space-type nuclear power systems have been considered. The typical characteristics that have been described represent an advanced nuclear power system that can be obtained with extrapolation of the present state-of-the-art. In the following discussions various components and problems of such a system are considered in greater detail.

Thermodynamics and Working Fluids

Working Fluids — the selection of a working fluid for a nuclear auxiliary power system depends on many factors. It is not within the scope of this discussion to evaluate in detail all the factors that bear in selecting a particular fluid. Factors that must be considered in fluid selection are as follows:

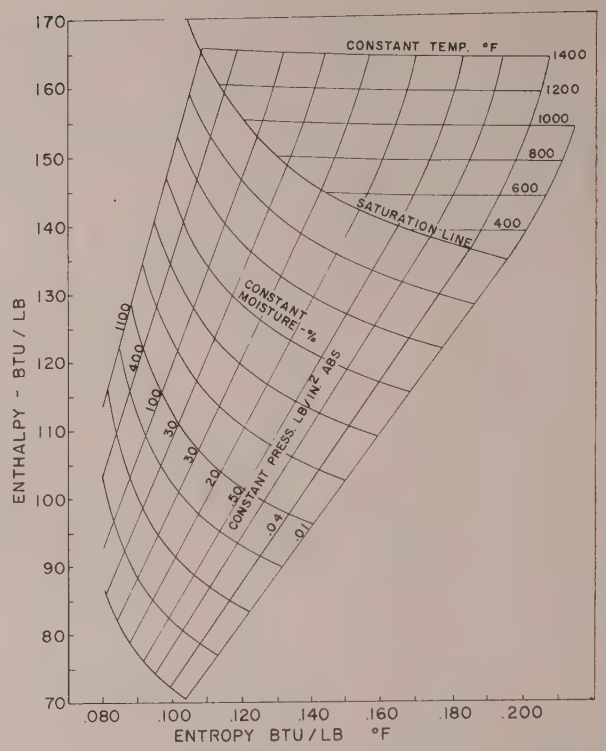


Fig. 14 — Total heat-entropy diagram for mercury (from "Thermodynamic Properties of Mercury Vapor," by L. A. Sheldon. ASME paper 49-A-30)

1. Neutronics: The fluid must be amenable to reactor use. That is, it should not have high neutron absorption properties, nor should it become activated due to neutron bombardment.

2. Thermodynamics: A satisfactory fluid must have thermodynamic properties such that interrelationship of pressure, volume, temperature, enthropy, and the like, combine to make a satisfactory power cycle. Fluid pressure should not be excessive at the necessary high temperatures, but it is also desirable that the pressures not be so low that very high specific volumes are involved. The phase characteristics (that is the forms and rates of formation of the fluids) are also important. For example, does the fluid exist in a solid state in storage, or, does it change molecular specie as it expands in the turbine?

3. Heat-Transfer Characteristics: Obviously, a desirable fluid must have good heat-transfer characteristics and not involve undue fluid friction loss.

4. Corrosion Properties: Materials behavior is an extremely important area, particularly at the high temperatures involved. The fluids must be compatible with reactor, turbine, pump, radiator, and other system components materials of construction.

5. Fluid Stability: The fluid must be stable at high temperatures, and must withstand the radiation environment of the reactor. Organic fluids are ruled out because they are not stable at the temperatures involved in the present systems.

6. Effect on Startup Requirements of the Systems: Possible problems in system startup may involve the fluid phases. If the fluid exists as a solid under normal environmental storage conditions, then it is

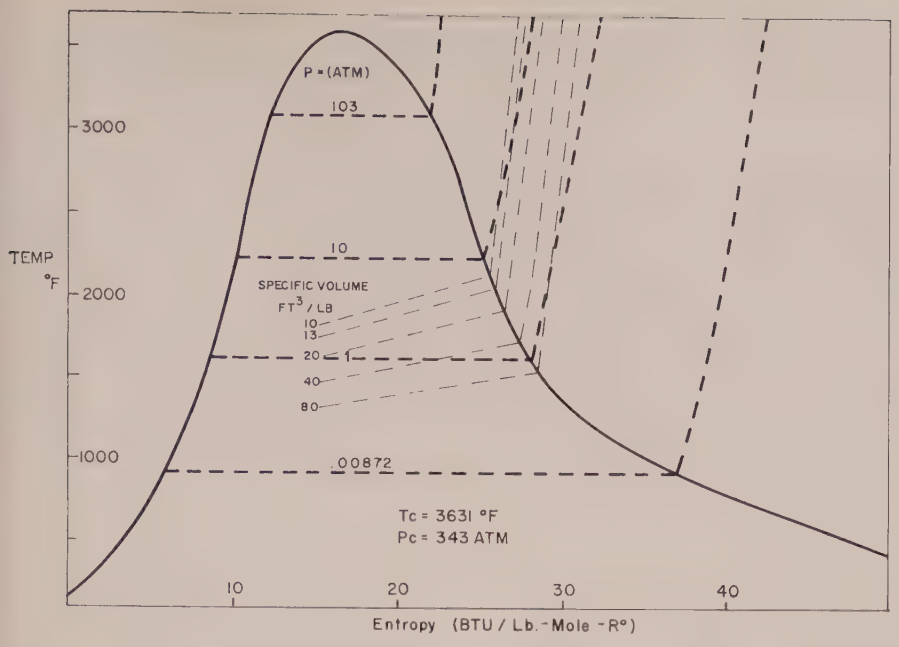


Fig. 15 — Temperature-entropy diagram for sodium (from NAA SR-62)

necessary to determine how to liquefy it and bring the system into operation.

7. State of Knowledge: If there is not sufficient available data on the fluid, the magnitude of the necessary data development program must be determined. In addition, the operating experience obtained with a given fluid must be considered.

8. Availability and Cost: The availability of the fluid and the expense of the necessary purification processes must be considered.

Secondary consideration must be given to other factors, such as: (1) diffusion characteristics of the fluid through containment materials, (2) possible logistic and handling hazards, and (3) effect on turbine and compressor design (sonic velocity in the fluids). Only after very extensive evaluations are performed can a fluid be selected. Moreover, the fluid cannot be selected independent of the thermodynamic cycle to be used.

Some Possible Fluids — Some fluids that can be considered for application to nuclear auxiliary power systems are: helium, nitrogen, mercury, sulfur, phosphorous, rubidium, and sodium. Helium and nitrogen behave very much as perfect gases, and there is a fair amount of data on their properties. A reasonable amount of data is also available on mercury and sodium. However, for sulfur, phosphorous, and rubidium data is limited. Some *T-S* diagrams, *P-V-T* diagrams, and phase diagrams for mercury and sulfur, phosphorous, rubidium, and sodium are given in Figs. 14–25. The data shown for sulfur, phosphorous, and rubidium have been synthesized from somewhat meager available information and are useful for preliminary evaluations only. The critical point shown for rubidium is very much in question. These diagrams, however, show the general characteristics of these fluids. Mercury, rubidium, and sodium behave in a somewhat normal fashion; that is, there are no radical phase changes throughout the temperature and pressure range shown. Sulfur and phosphorous are quite different. The peculiar shape of the dome for sulfur occurs because of the various molecular species that exist at different temperatures. The molecular composition

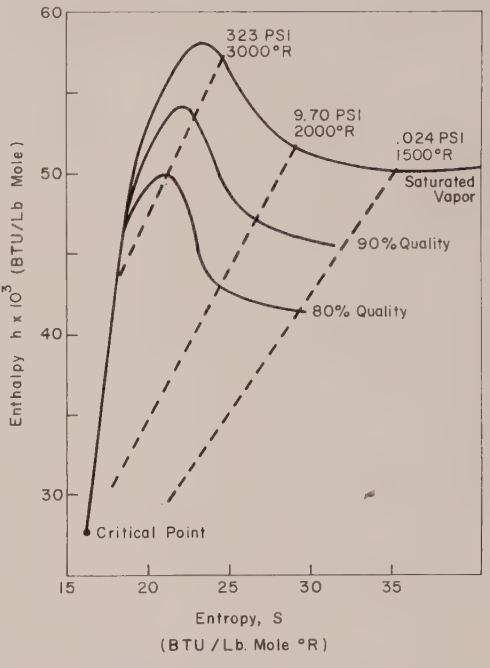


Fig. 16 — Mollier diagram for saturated and wet sodium vapor (from NAA SR-62)

of sulfur at various temperatures is shown in Fig. 25. Phosphorous similarly undergoes a phase change. It precipitates out as the solid red-*V* form at low temperature. This could present serious problems in power system machinery. However, as shown in Fig. 22, the time required for phosphorous to solidify is finite, and if the circuit times are short, perhaps litle or no precipitation would occur.

Because the above fluids are ones that might be considered for application to nuclear auxiliary power units, they have been used in the cycle calculations that are given in the next section.

Thermodynamic Cycles — A basic criteria concerned in selecting a thermodynamic cycle (and the working fluid) are overall minimum system weight and volume, good reliability, long lifetime, and a

reasonably short system development program. If we restrict ourselves to considering single working fluid systems, there are essentially two types of cycles that can be used; vapor cycles and gas cycles.

The vapor cycle is attractive for a number of reasons. From the thermodynamic point of view it is highly desirable to have most of heat added at a single high temperature, and to remove it at a single low temperature. This is accomplished with a fluid in two-phase equilibrium, because the fluid can be heated at constant pressure in a region under the vapor dome of the *T-S* diagram, where temperature depends only on pressure. Moreover, in a vapor cycle, the vapor can be expanded to extract work while liquid compression which requires relatively little work can be used to complete the cycle. For this reason, relatively inefficient pumps can be tolerated without appreciable loss in cycle efficiency. The classic Rankine cycle is illustrated in Fig. 12. Practically all the heat is added at constant temperature, and the only significant irreversability occurs in the expansion process. A difficulty that may be encountered in operation of a Rankine cycle is moisture condensation in the vapor stage damaging turbine blades. However, at low power levels, a single-stage turbine will be used, and the transit time through the turbine will be so short that moisture formation will be minimized. This problem requires further study. The problem of high turbine moisture can be resolved by modifications of the simple Rankine cycle using superheating or extraction. Generally, such modifications result in reduced cycle efficiencies because of reduced boiler

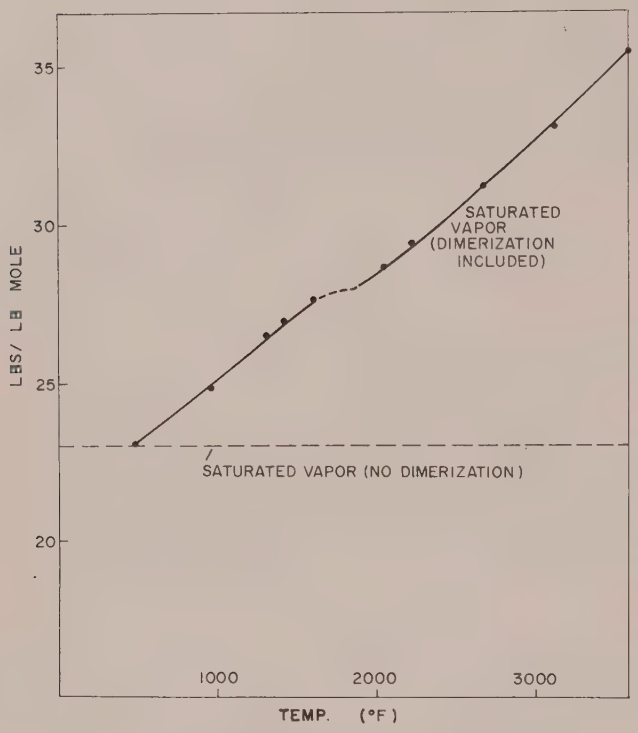


Fig. 17 — Pounds/pound mole for saturated sodium (from NAA SR-62)

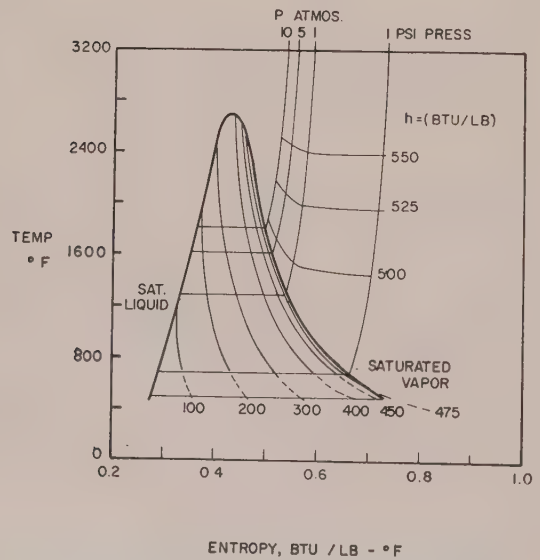


Fig. 18 — Temperature-entropy diagram for rubidium

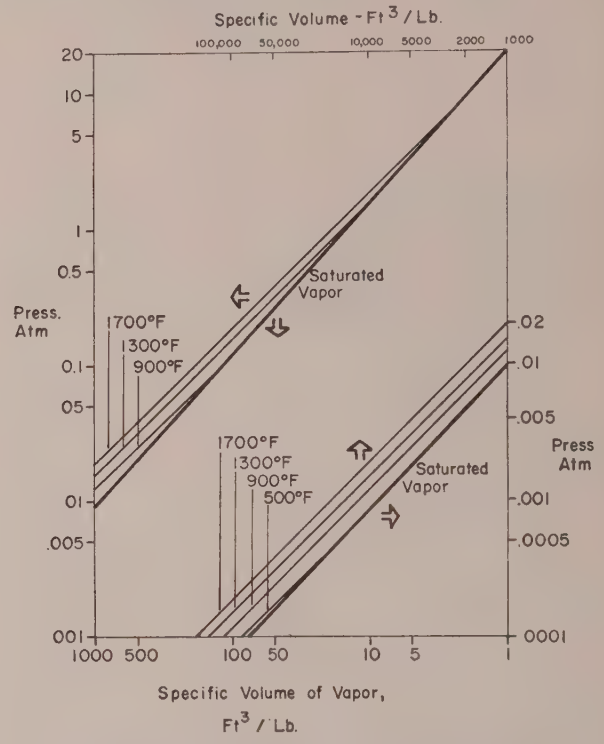


Fig. 19 — Pressure-volume curves for rubidium vapor

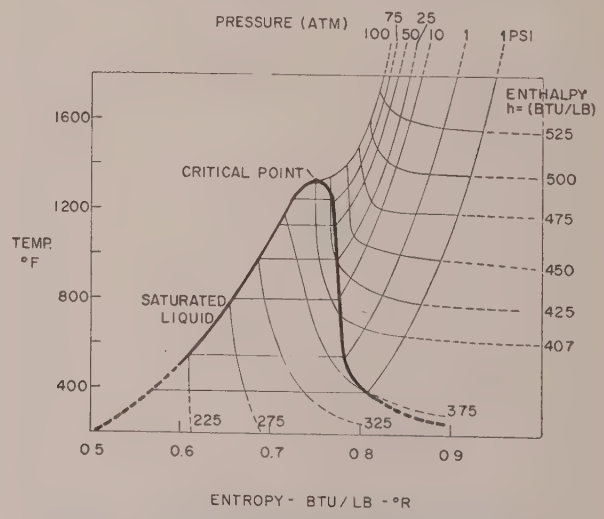


Fig. 20 — Temperature-entropy diagram for phosphorus (α white liquid – P_4 gaseous)

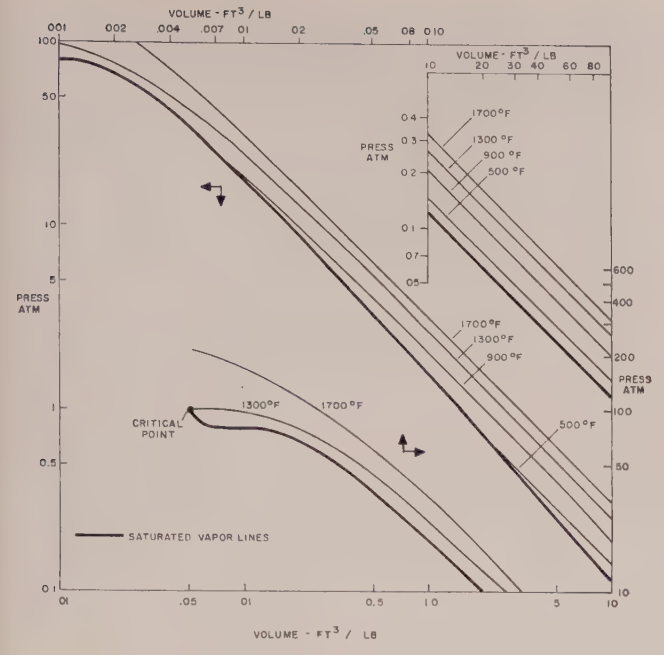


Fig. 21 — Pressure-volume curves for phosphorus (P_4 state gaseous) from $PV = Z\,RT$

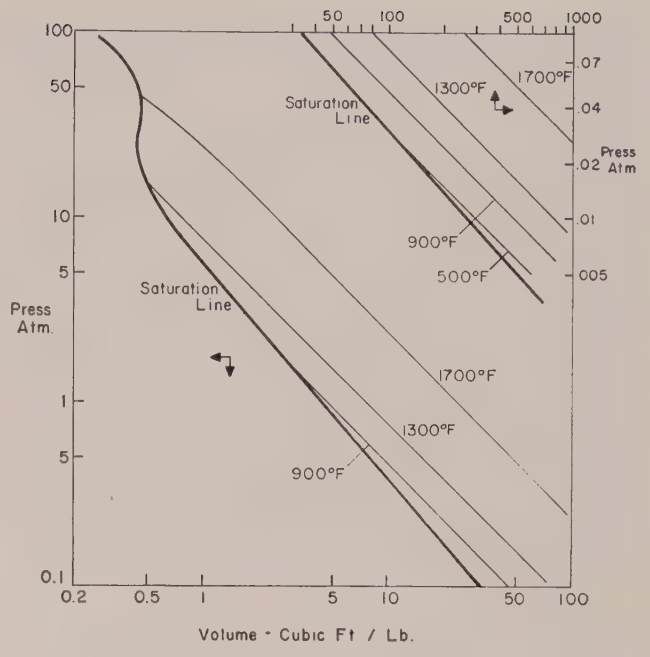


Fig. 24 — Pressure-volume curves for gaseous sulfur (S_2, S_6, S_8 mixture)

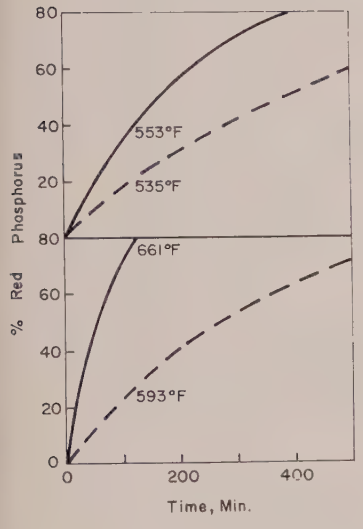


Fig. 22 — Conversion rate of α white to red V phosphorus (phosphorus as substitute for mercury in power cycle — John J. Grebe, Mon 344, July 1, 1947)

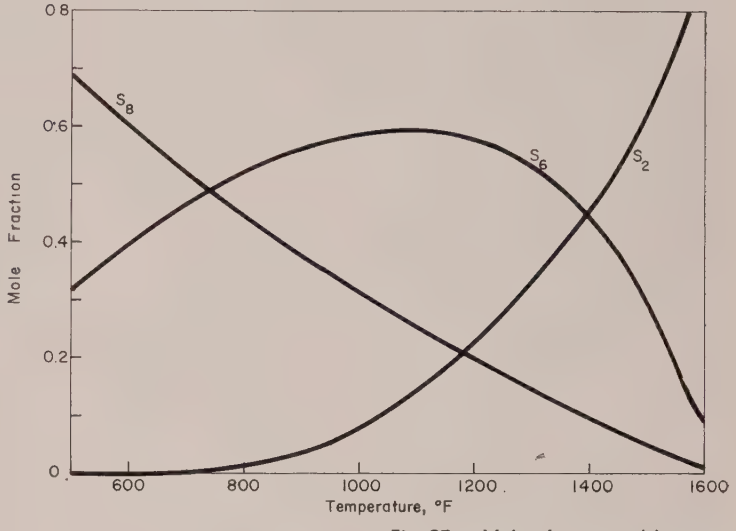


Fig. 25 — Molecular composition-temperature for saturated sulfur vapor

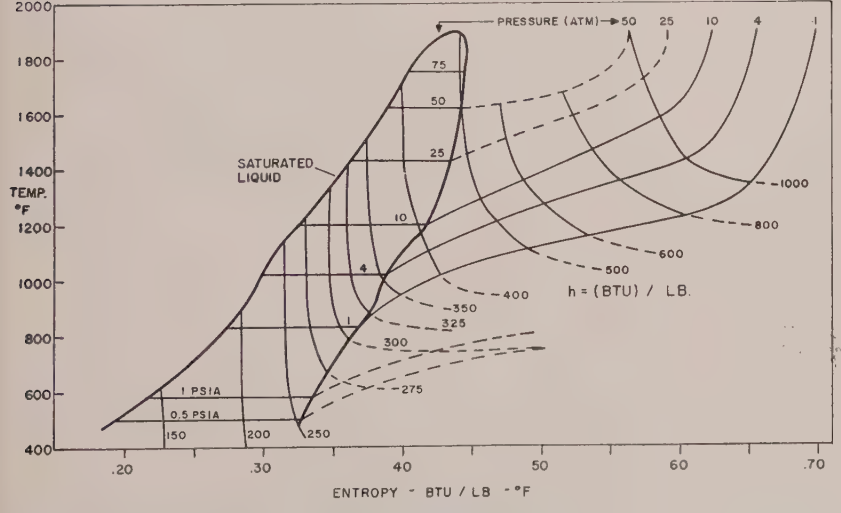


Fig. 23 — Temperature-entropy diagram for sulfur (S_8 in liquid; S_2, S_6, and S_8 in vapor) — chemical equilibrium state

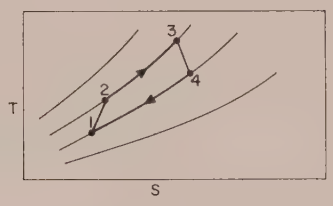
(reactor) temperature and increased complexity which are not justified because of weight and reliability difficulties.

Gas cycles offer attractive features because inert gases such as helium or nitrogen could be used, thereby assuring essentially no radioactivity problems due to neutron bombardment, greatly simplified corrosion problems, and no zero-gravity condensation problems. Unfortunately, the radiator areas required for gas cycles make them unattractive for the power levels and source temperatures of present interest. The chief disadvantage of the gas cycle lies in the fact that the back-work ratio is very high (of the order of 70% or more) and this irreversibility in the compressor affects the cycle performance. This results in large radiator size. In the simple Brayton gas cycle (Fig. 26A and 27A) a large part of the heat is added at low temperatures, tending to give low cycle efficiency. This can be partially compensated for by regeneration, but regeneration is possible only if the turbine outlet temperature exceeds the compressor outlet temperature. At low pressure ratios this may not be the case. At higher pressure ratios the irreversibilities in the turbine compressor tend to offset the gain due to regeneration, and there is a pressure ratio limit above which no net work results. This is especially true at high temperatures and low powers where turbine and compressor efficiencies tend to be relatively poor. There is also a low limit to the pressure ratio below which a nonregenerative Brayton cycle will not produce work.

Under certain circumstances the band of pressure ratios within which the regenerative cycle will be useful is very narrow. Brayton cycle modifications that extend the range of available pressure ratios and increase cycle efficiency include multistage compression and expansion, with several stages of intercooling and reheat as shown in Fig. 26B and 27B. However, even with these complex modifications, the radiator areas required are still much too large in comparison to vapor cycles. Other types of

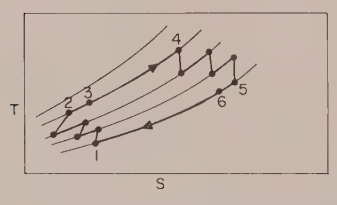
Table 3 — Comparison of Rankine and Brayton Cycles

Rankine Cycle	
Turbine efficiency	70%
Pump efficiency	65%
Alternator efficiency	85%
Condenser effectiveness	85%
Temperature difference, pump inlet-sink	100 F
Maximum fluid temperature	1,600 F
Negligible fluid friction pressure drop	
Brayton Cycle	
Turbine efficiency	80%
Compressor efficiency	70%
Alternator efficiency	100%
Regenerator effectiveness	85%
Temperature difference, compressor inlet-sink	100 F
Complete intercooling and reheating	
Equal expansion and compression ratios in each stage	
Regenerator and reactor pressure drops equal to 5% of static pressures	
Maximum fluid temperature	1940 F



1-2 Compression in the compressor
2-3 Heat addition in the reactor
3-4 Expansion in the turbine
4-1 Heat rejection in the radiator

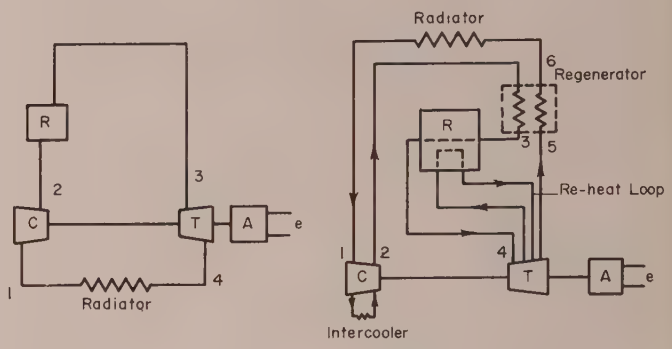
a) THE SIMPLE BRAYTON CYCLE



1-2 Compression with 2 stages of intercooling
2-3 Preheating in regenerator by cooling of exhaust from 5-6
4-5 Expansion with 2 stages of reheat
6-1 Heat rejection in the radiator

b) REHEAT, INTERCOOLING, REGENERATIVE BRAYTON CYCLE

Fig. 26 — The Brayton gas cycle



a) Simple Cycle

b) Re-heat, Regenerator, Intercooler Cycle

R-REACTOR T-TURBINE
C-COMPRESSOR A-ALTERNATOR

Fig. 27 — Brayton gas cycle configurations

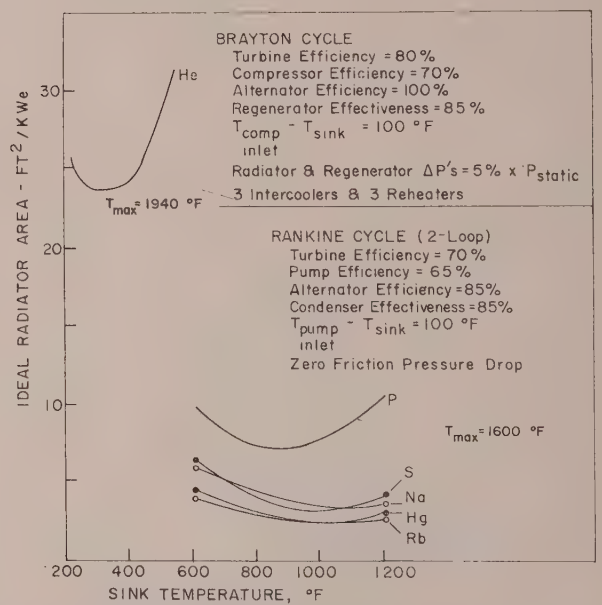


Fig. 28 — Comparison of ideal radiator areas for Rankine and Brayton cycles (spherical radiator, shape factor = 1.0, emissivity = 1.0)

gas cycles such as the Stirling cycle or hybrid engines can be devised, but basically they involve larger radiators than vapor cycles and/or greater complexity; therefore, possibly greater difficulty in obtaining reliability.

The simple two-loop Rankine vapor-cycle performance is compared in terms of ideal radiator area to the somewhat complicated regenerative-intercooling-reheat Brayton cycle performance (Fig. 28). Radiator areas required for the complex Brayton cycle are anywhere from 3 to 10 times larger than those for the Rankine cycle. The assumptions used in these calculations are given in Table 3.

Reactor and Shield

Reactor —

Size — As has been discussed previously, small size is extremely important in the design of a reactor for use in space. In the lower power applications (less than 1 mw), the reactor shield is a major source of weight, and because shield weight varies as the square of the reactor radius, the smallest possible reactor size should be used. It is for this reason that fast reactors appear best suited for the present applications. A fast reactor core can typically be of the order of 6 in. in diameter with a critical mass, in the range of 25–50 kg of U^{235}. Thermal or epithermal reactors may find application at high power levels, where the reactor size is dictated by heat-transfer considerations and not critically. Beryllium-oxide moderated units can be made as low as 22–24 in. in diameter, resulting in critical masses of 30 or more kg of U^{235}. Various possible metal hydride moderated designs could also be considered. These designs achieve diameters and critical masses considerably smaller than those of the beryllium-oxide design, although they do not reach the size that can be achieved by a fast reactor. Hydride-moderated reactors have a problem because of the possibility of losing hydrogen from the hydride at the high temperatures involved. Loss of hydrogen results in less of reactivity; for large losses the reactor might no longer maintain the chain reaction.

Zero-Gravity Operation — Operation in a zero-gravity field introduces a number of special problems in the reactor. These problems can be divided into two areas: startup and heat transfer. The startup problem requires careful design analysis for all possible conditions. The heat-transfer problem is discussed in greater detail in the next section.

Heat Transfer — If the coolant fluid is boiling in the reactor (two-phase operation), the zero-gravity problem is of extreme importance in regard to reactor heat transfer. There are at least three possible approaches to heat-transfer design for a reactor of this type, all three of which may be used in certain applications.

The first approach is to have the fluid go all the way from liquid to saturated vapor in the reactor. This eliminates the problem of a zero-gravity phase separator, and makes an additional recirculating pump unnecessary. It has the disadvantage of causing widely varying heat-transfer characteristics as the flow stream changes from laminar, to two-phase turbulent, to mainly vapor phase going through the reactor. This method appears to be feasible and is the simplest from the design standpoint. However, it can be used only at relatively low power levels, where a large heat-transfer surface is available, and the volumetric vapor flow is not unduely large. The method requires the greatest amount of heat-transfer area of any of the three methods considered here.

A second possible method of heat transfer is to have the fluid go from liquid to liquid-plus vapor in the reactor, and then separate the fluid into a vapor stream and a liquid stream in a zero-gravity separator. This method results in the highest allowable reactor heat fluxes because maximum heat-transfer coefficients are obtained for forced-convection boiling heat transfer. It has the disadvantage of the added complexity of the zero-gravity phase separator and possibly of an additional recirculating pump.

A third method of operation is simply to maintain the working fluid in the liquid state while in the reactor. In the case of a two-loop boiler-connected system, the reactor coolant could remain liquid at all times. In the case of a single loop system the liquid could be flashed to a lower temperature and pressure after leaving the reactor to obtain 5–15% vapor. This process is very inefficient for most fluids and means that the reactor must operate at higher temperatures and pressures than in the situation where the fluid boils. However, maintaining a liquid in the reactor at all times essentially eliminates the variation in percentage voids and for high power levels this mode of operation may be required.

High-Temperature Materials and Corrosion — High-temperature operation must be used to minimize the size of the radiator, and for high power levels, to reduce its large weight. High-temperature operation means using some type of refractory metal. Tantalum can be considered for a fast reactor, but for thermal reactors the high thermal neutron absorption cross-section makes it undesirable. Molybdenum or molybdenum with special coatings are excellent for use in high-temperature applications, and have good corrosion characteristics with most fluids. Development of welding and fabricating methods is required, especially for structurally complicated designs. There has been a considerable amount of development work on fabrication of molybdenum and it has good corrosion characteristics up to extremely high temperatures. Chromium is another good possibility, although it has a large number of fabrication problems and is extremely difficult to work. Corrosion behavior at these high temperatures must be evaluated for nearly all these materials, depending upon what fluid is used in the reactor. Corrosion probably limits the operating temperature, rather than materials strength.

Safety and Control — By maintaining the reactor shutdown until it is a large distance from earth or in orbit, safety probably will not become a major problem. With uranium-fueled reactors, the quantity of uranium is not a serious hazard. If plutonium is used in the design, alpha activity from the plutonium could be extremely hazardous if it were all released.

The control of the boiling reactor types considered in this application should be good, because the unit can have the inherent safety feature that increased boiling caused by increased temperature decreases fluid density in the unit which decreases the reactivity. A reactivity decrease will reverse the tendency to increase power that might have initially caused the higher temperature. However, this must be studied in detail to make certain that no insta-

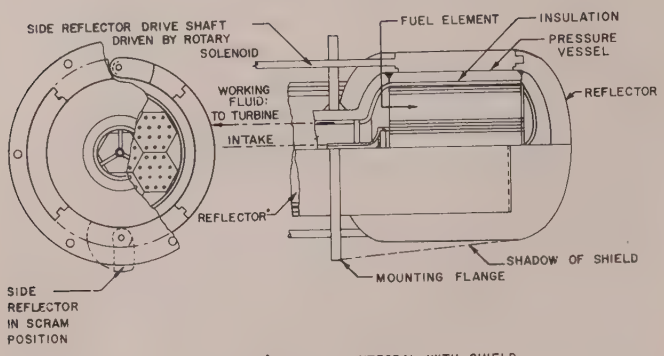


Fig. 29 — Schematic of fast reactor design

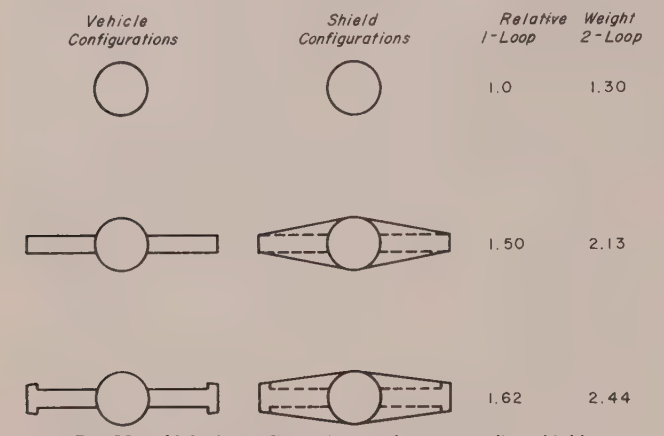


Fig. 30 — Vehicle configurations and corresponding shield configurations

bilities exist in the system.

Mechanical Design — Fig. 29 illustrates schematically the mechanical design of a possible fast reactor for space application. The fuel elements consist of hexagonal shapes with internal axial holes, through which the fluid passes. A pressure vessel surrounds the core, with the reflector outside of the pressure vessel. The unit is controlled by sliding or rotating portions of the reflector. The design shown is largely schematic, but gives some indication of what the unit might look like.

Shielding

Weight — It has already been pointed out that the shield contributes a large portion of the power-system weight. Two methods used to reduce the shield weight are: the use of a shadow shield (which reduces the surface that the shield covers), and separation of the power system from the vehicle (which reduces the radiation intensity at the vehicle). Separation reduces the shield cross-section required for a given shadow cross-section at the vehicle.

The influence of vehicle configuration on shield weight is an important factor. This effect is illustrated in Fig. 30. Adding protrusions to the vehicle drastically increases the shield weight. This is particularly true because the shield cross-section must not be re-entrant in order to prevent "second scattering" back to the vehicle. Close coordination of vehicle design and power-system design becomes an important factor in obtaining the overall best

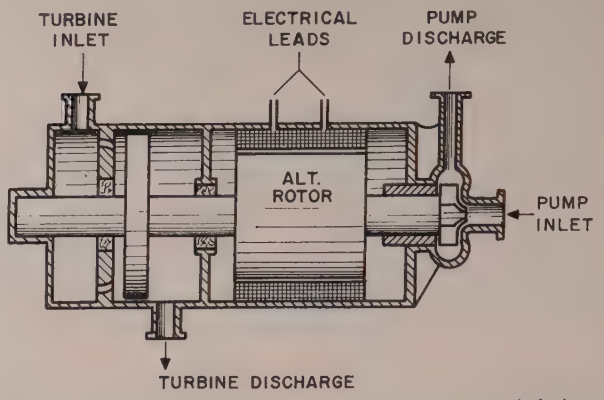


Fig. 31 — Schematic of turbine-alternator-pump canned design

lightweight system.

Materials — In general, the best type of shield will use some combination of neutron-shielding material which contains a large percentage of hydrogen- and gamma-shielding material which is of high density such as tungsten. The largest part of the tungsten will be placed as close as possible to the reactor. However, the shield must be laminated to reduce the capture-gamma dosage. For the neutron shield, zirconium hydride or lithium hydride may be considered. Boron also can be used to capture the neutrons that are slowed down, and thus reduce the amount of secondary gamma radiation produced in the shield. Considerable effort must be spent in the shield design, to optimize the use of materials and obtain as light a configuration as possible.

Radiation Scattering — Both neutrons and gamma rays can scatter from very small amounts of materials that are outside the shadow cone of the shield, and because the total radiation is large, it takes only a very small amount of material situated outside of the cone angle to scatter a large quantity of radiation back to the vehicle. Therefore, wings, antennas, radiator, and the like which might be extended must be maintained within the cone angle of the shield; or at least if they are extended beyond this cone angle, careful consideration must be given to scattering, to make certain it is not contributing a major dose to the vehicle.

Radiator and Condenser

Radiator — The radiator presents one of the major design problems in a nuclear auxiliary power system for space. At the low power levels (10–20 kw) radiator weight is not a major factor, but radiator volume is important. At higher powers radiator volume is of paramount importance, and in addition, radiator weight may constitute as much as 25% of the total power system weight, for say 1000-kw output systems.

An important factor in radiator design is selection of a configuration that provides for a maximum area utilization; that is, a design that permits the radiator surfaces to "see" as much space solid angle as possible. Shape factors for some plate-type, two-dimensional radiators are summarized in Table 4. Note that the shape factor for a cylindrical cross-section is only 0.5. This is because the inner surface "sees" only itself. The best shape factor, of course, is for the flat plate which completely "sees" space. Such a configuration may not be structurally desir-

Table 4 — Characteristics of Two-Dimensional, Plate-Type Radiators

Shape[a]	No. of Elements	Angle between Elements, deg	Shape Factor, S_f	Actual Area/Unit Axial Length	Effective Area/Unit Axial Length	Relative Weight/Unit Effective Area
A ○	-	-	0.500	$4\pi r$	$2.0\pi r$	1.000
B —	2	180	1.00	$4 r$	$4.0 r$	0.500
C ⅄	3	120	0.866	$6 r$	$5.2 r$	0.577
D +	4	90	0.707	$8 r$	$5.66 r$	0.705
E ✳	5	72	0.588	$10 r$	$5.88 r$	0.850
F ✳	6	60	0.50	$12 r$	$6.0 r$	1.00

[a] All shapes circumscribed by a circle of radius r.

able. The three-element configuration offers good shape factors and a much improved structural rigidity, and would, therefore, be a likely configuration.

Another very important consideration in radiator design is the effect of micro-meteorite damage. A very light design of minimum possible wall thickness will tend to be much more vulnerable to micro-meteorite penetration. If the radiator wall is thickened, then micro-meteorite vulnerability will be reduced at the expense of increased system weight. Micro-meteorite data at present is very meager, and will require very extensive study to yield the optimum radiator design which provides desired reliability at a minimum weight.

Another important factor in radiator design is coordination with the vehicle design. In a re-entrant type vehicle where the power system might have to be retracted to the vehicle, some provisions for storing the radiator system in the vehicle is required. In either re-entrant or non-re-entrant vehicles, the separation distance allowed will be determined by the power system weight and the maneuver rates required by the vehicles. The separation distances may or may not provide sufficient room to completely contain the radiator within the shadow cone. If the radiator cannot be contained within this cone, it must protrude out where it may "see" the reactor and scatter radiation back to the vehicle. As noted above, such a design must be very carefully evaluated to assure that any scattered radiation is not above excessive limits.

Rotating Machinery

The primary problems involved in the rotating machinery are in bearings, seals, corrosion, and high-temperature magnetic problems in the alternator. The problem of seals is most readily solved by the use of completely canned designs, in effect, eliminating seals to the outer environment. Such a design is illustrated in Fig. 31. The bearing problem must be resolved through developmental effort. Probably it will be desirable to use the working fluid as the bearing lubricant. This would eliminate the complexity of having another fluid, and also eliminate the possibility of contaminating the working fluid with lubricant.

The alternator design is complicated because vapors will be present in the alternator, and high temperatures are necessarily involved. A brushless magnetic switching-type alternator appears quite attractive for this application.

The condensate pump is directly connected to the turbine-alternator shaft, as illustrated in Fig. 31. The main problems for the pump are those of materials, fabrication, cavitation, and corrosion. As noted previously, moisture in the vapor coming to the turbine may present some corrosion-erosion problems in the turbine blade design.

Conclusions

General

1. Non-air-breathing auxiliary power systems are required over a very broad spectrum of power and duration. Chemical propellant and nuclear mechanical conversion systems are required for power levels greater than a few kilowatts. The chemical systems are best suited to short duration (several hours and less). Nuclear systems are the only possibility other than solar systems for long duration and high power.

2. The basic requirements of these systems are: minimum weight, minimum volume, maximum reliability, and compatibility with other components of the vehicle system.

Chemical Propellant-Mechanical Conversion Systems

1. These systems consist essentially of a gas generator and an energy-conversion device (usually a turbine) that extracts work from the gas after which the gas is ejected from the system.

2. Selecting a particular type of system is dictated largely by the energy sources (chemical fuels) available on the vehicle and the duration of the mission.

Nuclear-Mechanical Conversion Systems

1. Nuclear auxiliary power systems are essentially limited to application in space because of the necessity for low shield weight.

2. Even with the use of shadow shielding, the shield weight for 10–20 kw output systems is still quite important, and deserves careful consideration in design. In addition to the shadow shielding concept, shield weight is reduced by separation of the power system from the vehicle and by use of the smallest reactor design possible. (A fast reactor is attractive for this reason.)

3. The radiator is one of the most important components of the system in both volume and weight. High-temperature designs are required to obtain reasonably sized radiators. The radiator is, moreover, vulnerable to micrometeorite damage; the lightest design being most vulnerable. Overall system reliability is obtained at the expense of increased system weight.

4. The Rankine liquid-vapor condensing cycle appears more attractive than gas cycles because radiators about 1/3–1/10 the size are required for the Rankine cycle.

5. Corrosion and materials compatibility is an underlying problem in all of the system components. The system, in general, will be corrosion limited rather than strength limited.

6. Close coordination is required between vehicle and power-system design, to obtain the best overall minimum weight design.

Discussion of this paper will be found on pp. 128-129.

Many Unknowns in Development Of Nuclear Auxiliary Powerplants

— Dr. Robert H. Fox
Radiation Laboratory, University of California

IN GENERAL, this paper constitutes an excellent statement of state-of-the-art with respect to this relatively new class of powerplants. Most of the principal problems connected with the development of nuclear auxiliary powerplants for space application are discussed rather thoroughly.

A little more can be said about micro-meteorities. Although the data on micro-meteorite damage is indeed meager, there does exist information of the type which leads to design criteria for radiator materials. The absolute flux level is not well-known but the mass distribution law is known[a] and some experimental work on high-speed impact exists.[b] These can be used to show that the best radiator material is the one with lowest ρ^2/E (ρ = density, E = modulus of elasticity) at a given temperature. This minimizes the rate of penetrating hits for a given weight-to-power ratio for the radiator. Clearly, beryllium is superior to any other material in this respect at reasonable radiator temperatures. There remain problems of compatibility with the working fluids, fabrication, and permeability to working fluid in the vapor state.

In reference to safety and control of nuclear systems (p. 125), the point should be made that, depending on the nuclear properties of the working fluid, boiling reactors can have either a positive or negative dependence of reactivity upon boiling rate. Thus, one must be careful in considering the kinetics of this type of reactor.

The conclusions which are reached concerning the relative merits of gas and vapor cycles are certainly correct. Studies at this laboratory have indicated that a gas cycle would have to operate with a turbine inlet temperature more than 1000 F higher than a vapor cycle in order to compete on a radiator area basis. Consequently, the gas cycle does not avoid materials problems in the turbine, bearings, compressor, or reactor. At present, there seems to be little reason for giving it serious consideration for space applications.

With regard to the desirable properties for the working fluid, one further point should be made. Because of the stringent reliability requirements on space nuclear powerplants, the working fluid should be also a good lubricant. In this way, shaft seal problems are eliminated. Thus, possible working fluids must be examined for their chemical (corrosion), thermodynamic (cycle condition), and physical (lubrication) properties.

It is clear from this paper that there are numerous closely related problems to be solved in the development of large, lightweight nuclear powerplants. Much research and development work in a variety of technical areas remains before they become reality.

Direct-Conversion Units Have No Heat Rejection Problem

— H. D. Else
Westinghouse Electric Corp.

THE AUTHORS of this paper are interested primarily in arriving at a suitable system for obtaining electrical power at a relatively high level, 10 kw and greater, in an outer space environment. They review energy sources and conversion means and conclude that for the immediate future a thermal cycle, preferably a vapor cycle, with a heat engine and a somewhat conventional generator offers the greatest hope. For flights of short duration, they recommend the use of a chemical energy source and, for longer flights, a nuclear-fission source.

The suggested temperature boundaries are 1600 and 1000 F, 1600 F is not unreasonable for the turbine. In fact, it might be desirable to add superheat and raise this value. The heat rejection temperature (1000 F) is dictated by the condenser. The rather low values of condenser area they suggest result from this selection. Any revision downward increases the condenser area as a fourth power function of the absolute temperature.

Unless mechanical refrigeration is provided, the temperatures of all components having internal losses will stabilize at values above that established for the condenser itself. High-temperature restrictions on bearing lubrication are mentioned in the paper. With journal-type bearings axial location is particularly difficult to obtain. A hermetic construction exposes the generator windings to the operating fluid. A compatible material with good electrical conductivity may not exist. The reaction of mercury, for instance, with copper, silver, or even gold is well-known.

Insulation would be of ceramic-like materials which, as a class, are brittle and have thermal dimensional coefficients very different from those of the wire to which they might be applied. Junctures of any dissimilar materials are subject to the same consideration. And, finally, the availability of a suitable magnetic material must be examined. With allowance for thermal gradients the temperature of the core will be in the region of the Curie points of commercial lamination steels.

An incidental complication is the effect of high-speed, rotating apparatus on the guidance control of the vehicle.

With these considerations in mind, one cannot assume that the dimensions, weights, and efficiencies of the counterparts that operate within the atmosphere are applicable to apparatus for use in space.

The appeal of direct-conversion devices is not limited to the hope for greater reliability resulting from the elimination of moving parts. The higher goal is the creation of a system that does not evolve quantities of unusable heat which must be rejected. The "fuel" in a "fuel cell" combines to produce electricity with very little resultant heat. A solar generator which accepts only frequencies to which it is responsive would presumably have only its ohmic losses to dissipate. Even though all frequencies are accepted, the result is not too different provided the response is in the region of maximum energy of the solar spectrum. It is quite possible that a breakthrough in direct conversion from energy or heat will occur before suitable materials and mechanisms for a space type, heat-engine cycle are produced.

Thermal Stress Is Serious Limitation On Reactor Power Output

— A. P. Frass
Oak Ridge National Laboratory

IT IS NICE to see a comprehensive unclassified discussion of the satellite powerplant problem, especially when one finds himself in quite general agreement with the overall conclusions reached on such a complex collection of problems. There are some points wherein my point of view differs somewhat from that of the authors' and some comment seems to be in order. One of these has to do with the effects of neutron spectrum on reactor core size. It is stated that a fast reactor will be substantially smaller than a thermal or epithermal reactor. While most thermal reac-

[a] "Incidence of Meteor Particles upon the Earth," by A. A. Weiss. *Australian Journal of Physics*, Vol. 10, September, 1957, pp. 397–411.
[b] "Preliminary Investigation of High Speed Impact: Penetration of Small Spheres into Thick Copper Targets," by A. C. Charters and T. S. Locke, Jr., NACA Research Memo A58 D26, May 28, 1958.

tors are large and most fast reactors are small, it is perfectly possible to build essentially thermal reactors in very small sizes. The Los Alamos water boiler, for example, required only 1 kg of U^{235} and was contained in a sphere only about 1 ft in diameter. There is no reason why even smaller thermal reactors cannot be built, particularly if a thick reflector is used. I feel that the reactor core size is more likely to be determined by heat removal considerations, the effects of which are discussed by the authors.

While it was not mentioned, in the writer's experience perhaps the most important limitation on reactor power output is that of thermal stresses in the reactor core structure, particularly in the fuel elements. If a liquid metal is used as a reactor coolant, heat fluxes of the order of millions of Btu/hr-ft^2 can be obtained. These high heat-transfer rates lead to steep temperature gradients in fuel elements, large internal temperature differences, and hence severe thermal stresses. As a reactor of this type is cycled from low to high powers, these thermal stresses fluctuate with the power even if the mean reactor operating temperature is held constant. The resulting thermal strain cycling may lead to cracking in structural elements in the reactor core. This is likely to be an even more important limitation on reactor power density than those posed by either nuclear or heat transfer considerations.

The writer agrees that a thermodynamic cycle working fluid should be chosen to permit operation at high temperatures with moderate pressures. A radiator weight optimization study such as that summarized depends heavily on the minimum acceptable thickness of the radiator tube wall assumed for the analysis. An important consideration which will really determine the minimum allowable radiator wall thickness is that of leak-tightness. It will be essential to obtain an absolutely leak-tight system so that the working fluid will not be lost from the system over a period of many months or even a few years. The difficulties associated with achieving a vacuum-tight system cannot be appreciated except through bitter experience. It is possible to build such systems and at Oak Ridge we have built them in sizes ranging from very small pilot plants to the large gaseous diffusion plants, but this can be done only if every detail of the design is carefully chosen to facilitate the achievement of vacuum-tightness. If the system is to operate at temperatures above 400 F it should be completely welded up, the wall thickness ought to be not less than 0.030 in., no castings should be used, all of the containing envelopes should be of wrought material with the direction of working parallel to the surface, and all welded joints should be designed to minimize weld cracking and thermal strain cycling effects. I should like to suggest that anyone seriously interested in building radiators suitable for satellite powerplants begin immediately by building some very simple radiator elements and try to operate them at high temperatures over extended periods of time and see if he can keep them completely vacuum-tight.

The matter of vacuum-tightness raises a final question. It is stated that a dual-fluid system will give increased reliability by minimizing the effects of fission product contamination of the primary reactor coolant. I wonder whether this shortcoming of the single-fluid system may not be more than balanced by the reduction in reliability of the dual fluid system as a result of leak-tightness problems in the intermediate heat exchanger and the complication presented by the extra pump drive, expansion tank, and related control equipment required for the dual-fluid system.

Lightweight Solar-Electric Systems Good for Space Applications

— William C. Cooley
National Aeronautics and Space Administration

SINCE THE authors so convincingly show that nuclear auxiliary powerplants are essentially confined in application to outer space, where a shadow shield can be employed, it would have been interesting if they had included a discussion of the space application of *solar* auxiliary powerplants for comparison. Assuming reasonable advances in the development of lightweight solar mirrors, one can estimate that the solar-electric systems may weigh no more than an essentially unshielded nuclear-electric system. If so, the solar systems become very desirable, particularly for manned space vehicles, in order to avoid the weight penalty of reactor shielding and the hazards associated with radioactive fission products. However, the title of this session is "Auxiliary Powerplants for High-Speed, High-Altitude *Aircraft*," and I can understand the authors reluctance to embark on a discussion of the aerodynamics of solar collectors.

ORAL DISCUSSION

— Reported by J. D. Rogers
Westinghouse Electric Corp.

Robert Curran, Boeing Aircraft Corp.: The use of a shallow nuclear radiation shield means that an additional power source would be required for use in the atmosphere (for manned vehicles). Another problem is that the temperatures would be too high during takeoff and re-entry for satisfactory operation of the radiator. One factor which must be remembered is all the power generated goes into electrical used within the vehicle and the waste heat from this equipment must be dissipated. Cooling for this purpose must be obtained at a penalty. Much better performance can be obtained from chemical fuels than is shown on the charts in the paper. By use of cyrogenic fuels with large cooling capacities, sfc's of 1.5 or even lower can be obtained. The cost of $500 to put a pound into orbit is close to that assumed by Boeing. A lot of rocket fuel can be bought, however, for the million dollars that a nuclear reactor might cost. Reactor cost must be considered in determining the cross-over point. In summation, chemical fuels may be much better than shown in the chart, but nuclear system would nonetheless be best for high power outputs used for extensive time periods.

Mr. Taylor: There may be a difficult starting problem with liquid propellants.

R. S. Hall, Allison Division, GMC: The Phillips-Sterling air-engine cycle may be very suitable and could give efficiencies up to 40%.

Dr. Cochran: Thermal efficiency alone is not the criteria. What is needed is the most simple, most reliable, and lightest weight unit. The Brayton cycle can also become attractive when the complication of regeneration is added. I concur that all energy produced must be dissipated in space. One possibility might be refrigerating certain critical compartments.

J. D. Rogers, Westinghouse Electric Corp.: What about the problem of meteorite damage, particularly to radiators and condensers?

Dr. Cochran: The reliability can be improved by increasing radiator weight or by providing for bypass of damaged portions of the system. There is a question of trade-off of weight and reliability, with reliability particularly important on manned vehicles. By increasing the radiator weight by a factor 2, the damage problem from meteorites would be decreased many fold.

the effect of

Residual Stresses

Strain-Peening

R. L. Mattson and J. G. Roberts
General Motors Corp.

THE PURPOSE of this experiment was to determine the role of residual stresses in fatigue strength independent of other factors usually involved when residual stresses are introduced. It consisted of an investigation of the influence of residual stresses introduced by shotpeening on the fatigue strength of steel (Rockwell C hardness 48) in unidirectional bending. Residual stresses were varied by peening under various conditions of applied strain. This process introduced substantially the same amount and kind of surface cold working with residual stresses varying over a wide range of values.

It was found that shotpeening of steel of this hardness is beneficial primarily because of the nature of the macro-residual-stresses introduced by the process. There is no gain attributable to "strain-hardening" for this material. An effort was made to explain the results on the basis of three failure criteria: distortion energy, maximum shear stress, and maximum stress.*

IN THE STUDY of fatigue strength of materials, the experimental determination of the role of residual stresses as compared with load stresses has always been difficult, because generally any method of introducing residual stresses also introduces other changes which independently might affect fatigue strength[1] (for example, changes in structure). It was the objective of this experiment to separate residual stresses from these other factors, and to study their significance to fatigue strength. The basis for the experiment is strain-peening[2] and follows a suggestion made by J. O. Almen some years ago.[3] This technique permitted the introduction of a wide range of macro-residual-stresses near the surface of specimens by surface cold working in such a way that the "strain-hardening" remained substantially constant. Fatigue tests and residual stress analyses were made of several groups of specimens and these showed good correlation between residual stress near the surface and the endurance

* Paper presented at meeting of Division 4 of the Iron and Steel Technical Committee, Detroit, Jan. 15, 1959.

[1] "Fatigue," by T. J. Dolan, B. J. Lazan, and O. J. Horger, p. 25. Pub. by American Society of Metals, Cleveland, 1954.

[2] U. S. Patent 2,608,752.

[3] "Fatigue Weakness of Surfaces," by J. O. Almen. *Product Engineering*, Vol. 21, November, 1950, p. 125. Also, private communications.

Induced by

upon Fatigue Strength

limits. The experiment also served to clarify questions about the relative roles of strain-hardening and residual stresses in the improvements in fatigue strength obtained by local cold-working operations such as shotpeening and surface rolling.

Description of Experiment

Material — Specimens were machined from hot-rolled SAE 5160 spring steel of 0.215 in. thickness and 1¾ in. width.

Machining and Heat-Treatment — Fatigue test specimens were prepared as follows:

1. Material was normalized (1675 F for ½ hr).
2. Specimens were machined to dimensions shown by Fig. 1 (equal amounts removed from opposite faces by shaping and grinding).
3. Austenitized (15 min in rectified salt at 1600 F).
4. Quenched in oil.
5. Pretempered (480 F for 1 hr).
6. Tempered (770 F for 2 hr between clamped plates and aircooled).

The resulting surface hardness, as measured by Rockwell C scale, was 48 ± 2. The resulting martensitic microstructure had some ferrite grains extending in places to a depth of 0.0004 in.

Presetting — After heat-treatment and prior to shotpeening, the specimens were given a conventional (in the spring industry) presetting operation. This was accomplished by means of curvature blocks and a hydraulic press. The curvature of these blocks was such that the curvature change from the free, flat condition (as heat-treated) to the preset curvature was sufficient to stress the central 6 in. of length to a nominal surface stress of 220 ksi (psi in thousands) which is slightly above the yield strength. This maximum stress was held for 0.5 min. All presetting was in the same direction as the subsequently applied fatigue stress.

Shotpeening — The strain-peening process is illustrated by Fig. 2. To the left, compressive strain-peening is shown. Here the specimen is held under

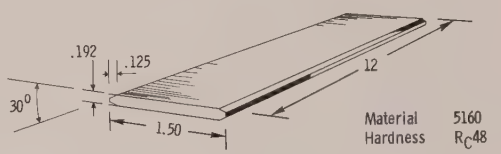


Fig. 1 — Leaf-spring fatigue test specimen

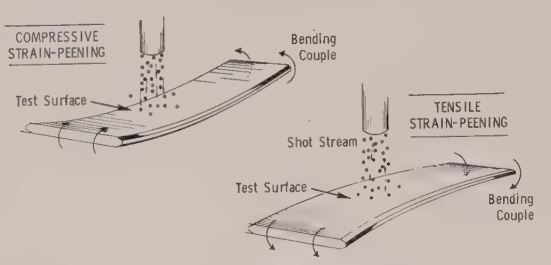


Fig. 2 — Illustration of strain-peening process

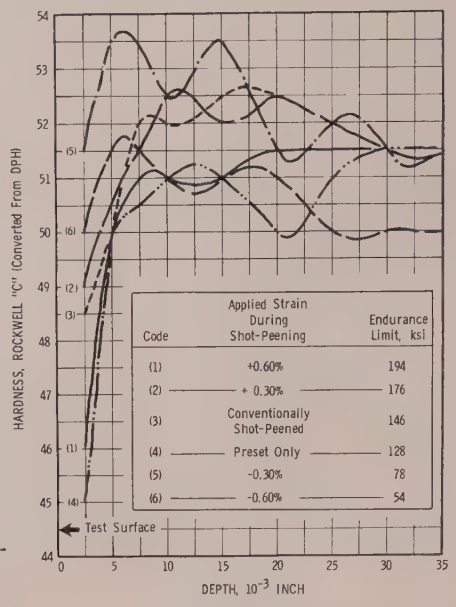


Fig. 3 — Hardness-depth distributions

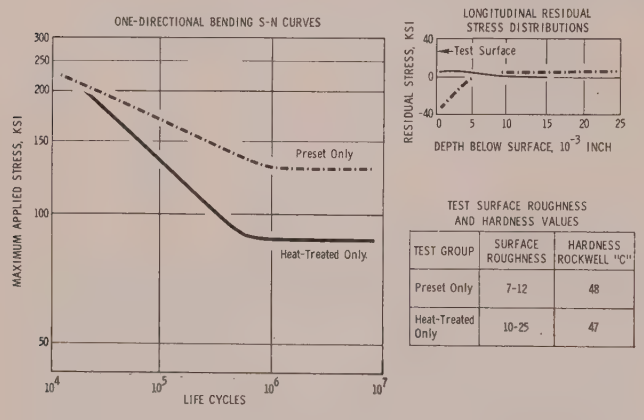

ONE-DIRECTIONAL BENDING S-N CURVES
MAXIMUM APPLIED STRESS, KSI
LIFE CYCLES
Preset Only
Heat-Treated Only
LONGITUDINAL RESIDUAL STRESS DISTRIBUTIONS
RESIDUAL STRESS, KSI
Test Surface
DEPTH BELOW SURFACE, 10^{-3} INCH
TEST SURFACE ROUGHNESS AND HARDNESS VALUES
TEST GROUP | SURFACE ROUGHNESS | HARDNESS ROCKWELL "C"
Preset Only | 7-12 | 48
Heat-Treated Only | 10-25 | 47

Fig. 4 — Influence of presetting on bending fatigue strength

a condition of longitudinal bending to a fixed curvature with the test surface subjected to a compressive strain during the actual shotpeening operation. At the right, tensile strain-peening is illustrated.

During peening, severe plastic flow occurs in a thin surface layer, upsetting the elastic strains due to the initial bending couple. The material beneath the thin surface layer remains unaffected by the peening. Upon release of the specimens from its holding fixture, relaxation occurs. The amount and direction of the relaxation depends upon the applied strain during peening. Hence, the magnitude of the resulting residual stress becomes a function of the applied strain during peening. For example, under conditions of tensile strain-peening higher residual compressive stresses are introduced than peening under a "free" condition. Furthermore, tensile stresses can be introduced by shotpeening under compression if the applied compressive strain is sufficiently high.

The shotpeening was done in an air-type machine designed especially for experimental work. Four specimens were held in a fixture which was rotated and traversed beneath a fixed shot-nozzle. A single shotpeening treatment was used for all peening done in this experiment. SAE 230 chilled iron shot, 0.006 C Almen intensity, and full visual coverage was used. Only the "test surface" was peened (the surface later subjected to tensile strains during fatigue testing).

Specimen Groups — The specimen groups studied are listed in Table 1. The "as heat-treated" group was included for reference and comparison with the "preset only" group. The "preset only" group serves as a baseline, since all shotpeened groups were preset prior to shotpeening. The various shotpeened groups have the same amount of work hardening but differ in residual stress. The final group was included to obtain some measure of the influence of surface roughness as produced by shotpeening on the fatigue strength.

All fatigue testing was done without any additional surface treatment.

The longitudinal curvature of each specimen was measured with an arc-height gage (sensitivity 0.0001 in.$^{-1}$, curvature) after heat-treatment, after presetting, and after shotpeening. Curvature that causes tension stress on the test surface was taken as positive. Average specimen curvature after strain-peening was dependent upon the magnitude and sign of the applied strain during peening. Curvature data are given in Table 2.

Photographs of the presetting equipment and of the curvature gage were presented in a previous paper.[4]

Metallurgical Examination — Hardness-depth curves in the vicinity of the test surface were made on individual specimens in order to evaluate the extent of strain-hardening. Hardness was determined by means of a Tukon tester with 1000-gm load and a diamond pyramid indentor. Results are shown in Fig. 3.

Fatigue Testing — One-directional bending type fatigue testing was used. All testing was done on constant-deflection type machines, operating at either 480 or 720 load cpm. Surface stresses were calculated from longitudinal curvature changes and the flexure formula. These were verified by both static and dynamic strain gage (electrical resistance type) measurements. Load stress was applied by means of a curvature block attached to a movable beam which was driven in fixed vertical motion. The other end of the beam was pin-connected to a fixed lower member on which two end blocks were attached, which support the ends of the specimen. Transverse stresses caused by the loading cycle were approximately zero over the entire test region as determined by resistance-type strain gages assuming a Poisson's ratio of 0.3. After an initial 10,000 stress cycles, each specimen was removed from the fatigue test machine, and longitudinal curvature was measured and compared to the initial curvature. Any appreciable change in longitudinal residual stress caused by these stress cycles would be detected in this manner. At least eight specimens per group were used to determine the S-N curve.

Residual Stress Measurement — Macro-residual-stresses were measured using the "beam dissection"

[4] "Effect of Shot Peening Variables and Residual Stresses on Fatigue Life of Leaf Spring Specimens," by R. L. Mattson and W. S. Coleman, Jr. SAE Transactions, Vol. 62, 1954, pp. 546–556.

technique developed in our laboratory. Essentially, this technique consists of analyzing a small rectangular cross-section beam segment of the fatigue specimen by careful removal of a series of thin layers and measuring the changes in longitudinal curvature. A careful grinding technique was used to remove the thin layers. The layers ranged in thickness from 0.001 to 0.020 in. with the thinner layers removed in the most important regions.

A more complete description of the technique is given elsewhere.[5, 6] This theory considers stresses in one direction only. Stresses perpendicular to the surface were taken to be zero.

The segments parted out from the fatigue specimens for residual stress analysis were $0.120 \times 0.192 \times 3$ in. for longitudinal stresses and $0.060 \times 0.192 \times 1.5$ in. for transverse stresses.

The residual stress distribution given in subsequent figures are those found in the individual segments. It is estimated that the standard deviation for such residual stress is of the order of 10 ksi, and for a thin surface layer, say 0.004 in., it may be greater.

One of the limitations of the techniques used is the lack of absolute measurements of stresses at the surface because of the removal of layers of finite thickness. Values referred to as being "near the surface" are actually those obtained from our measurement at a point of 0.001 in. below the surface.

Results and Discussion

Effect of Presetting — Since all of the specimens were preset prior to shotpeening, consideration will first be given to the influence of this operation on fatigue strength and residual stresses. The effect of presetting on residual stresses can be seen in the upper right diagram of Fig. 4, which shows residual stress versus depth for the first 0.025 in. from the test surface. It can be seen that presetting introduces about 35-ksi compressive stress near the surface and the "as heat-treated" specimen has a residual stress in this same region of about 5-ksi tension. The improvement in fatigue strength caused by presetting is shown on the left. Surface roughness and hardness were about the same in both cases. This improvement can be attributed to the more favorable residual stress distribution. However, strain-hardening cannot be ruled out completely. "Preset only" specimens will serve as our baseline for subsequent discussion.

Effect of Strain-Peening — The effect of variable residual stresses and constant strain-hardening (shotpeened specimens) on fatigue properties is illustrated by Fig. 5. To the left are S-N curves for six categories of specimens, and to the right are corresponding longitudinal residual stress distributions. The endurance limit of the preset specimen is approximately 128 ksi and its resdiual stress near the surface is about 35-ksi compression. If specimens of this type are conventionally shotpeened, the expected residual compressive stress, which is greater than that from presetting, is introduced, and an associated increase in the endurance limit may be noted. Still higher residual compressive stresses are introduced if specimens are strain-peened in tension. This results in a still higher endurance limit (approximately 194 ksi for the +0.60% strain-peening and 176 ksi for the +0.30% strain-peening). While the benefits of tensile strain-peening have been known for many years, this process has not been used extensively in practice. The present study is the first time, to our knowledge, that strain-peening under conditions of compression has been tried. If specimens are subjected to compressive strain during peening, residual tensile stresses may be introduced near the surface. For a strain-peening of 0.60% compression the peak residual stress in the affected surface layer is approximately 60-ksi tension and the endurance limit drops to approximately 55 ksi. This is a drastic drop. An intermediate result is obtained with −0.30% strain-peening. It is reasonable to assume that the effects of strain-hardening on all five shotpeening groups were about the same, since the total local strain caused by the impact of shot particles is obviously many times greater than the applied elastic strain

Table 1 — Specimen Groups Studied

As heat-treated
Preset only
Conventional shot peened
Shot peened under strain +0.60% (180 ksi)
Shot peened under strain +0.30% (90 ksi)
Shot peened under strain −0.30% (−90 ksi)
Shot peened under strain −0.60% (−180 ksi)
Shot peened free and stress-relieved for 2½ hr at 650 F.

Table 2 — Longitudinal Curvature Caused by Presetting and Strain-Peening

Specimen State	Average Specimen Curvature, in.$^{-1}$
After heat-treatment	0.0000
After presetting	+0.0010
After strain-peening (+0.60%)	+0.0145
After strain-peening (+0.30%)	+0.0100
After shotpeening (conventional)	+0.0062
After strain-peening (−0.30%)	+0.0005
After strain-peening (−0.60%)	−0.0060

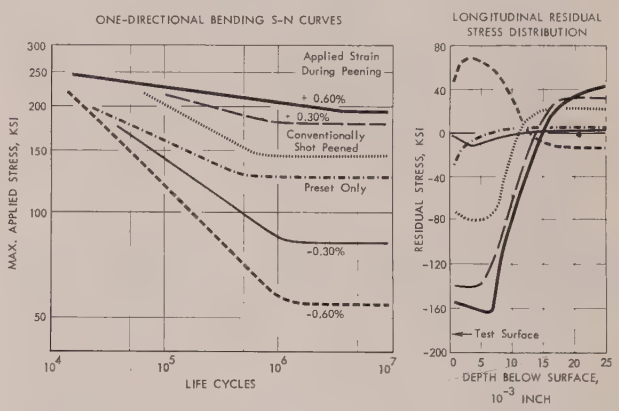


Fig. 5 — *S-N* curves and residual stresses for variously shotpeened leaf-springs

[5] "Theory and Procedure for Residual Stress Analysis as Practiced by Engineering Mechanics Department," by J. L. Rembowski. Research Laboratroies, General Motors Corp.

[6] "Derivation of Fundamental Equation of Residual Stress Analysis by Dissection," by L. G. Johnson. *Industrial Mathematics*, Vol. 5, 1955, pp. 7–13.

during strain-peening. Yet the endurance limits are drastically different. The primary differences observed are found in the residual stress distributions shown on the right-hand diagram. This is considered to be the most important observation made in this study. If strain-hardening were the more dominant feature of the shotpeening process, the endurance limits would have been more nearly the same.

A conventional AM (alternating stress versus mean stress) plot of these fatigue data would show how illogical it is to consider fatigue test results of this kind without considering the residual stress state. Such a plot would have a positive slope and go through the origin. It is quite clear that as the mean stress goes to zero the varying stress would also have to go to zero which, of course, is completely incorrect. However, if the residual stress is taken into account, and, as a first approximation consideration is given only to stress near the surface, a much more reasonable line would be obtained.

In Fig. 6 the endurance limit is plotted against residual stress near the test surface. For all of the specimens which were shotpeened, it can be seen that there is an approximately linear relationship. In addition, the "preset only" specimen actually lies to the right of the curve, indicating that for the same level of residual stress the endurance limit is somewhat higher. This is attributed to the smoother surface, and/or a more beneficial type of strain-hardening from the presetting operation.

Fig. 7 illustrates the influence of surface roughness. Conventionally peened specimens were given a stress-relief treatment of 2½ hr at 650 F and then fatigue tested, giving the results shown. The stress-relief treatment did not alter the hardness significantly, but the endurance limit dropped from 146 to 92 ksi. Residual stress analyses showed that the shotpeened, stress-relieved specimens had about the same residual stress near the surface as the "preset only" group (that is, – 35 ksi), as can be seen in Fig. 7. Yet the "shotpeened, stress-relieved" group gave a much lower endurance limit which primarily can be attributed to the increased surface roughness. When the "shotpeened, stress-relieved" group is compared with the group peened at an applied strain of – 0.30%, which has similar surface roughness, the difference can probably be accounted for by the more favorable residual stress in the "shotpeened, stress-relieved" group. A further comparison can be made between the "shotpeened, stress-relieved group" and the "as heat-treated" group. Here, in spite of increased surface roughness, the "shotpeened, stress-relieved" group has a higher endurance limit than the "as heat-treated" group. However, the "shot-

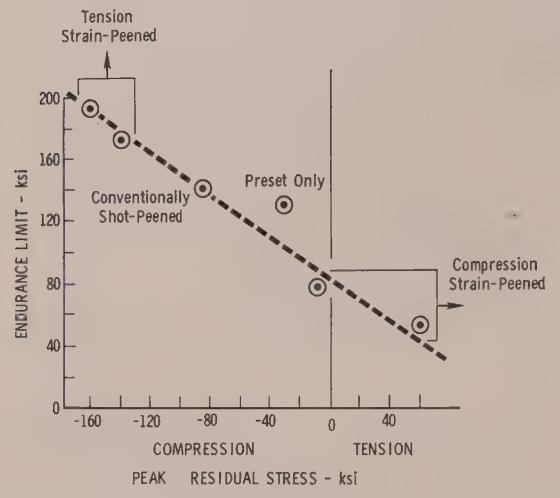
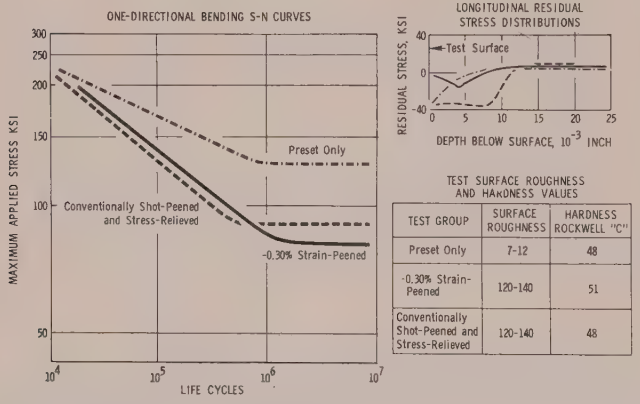


Fig. 6 — One-directional bending endurance limit as a function of peak residual stress near the test surface



TEST SURFACE ROUGHNESS AND HARDNESS VALUES

TEST GROUP	SURFACE ROUGHNESS	HARDNESS ROCKWELL "C"
Preset Only	7-12	48
-0.30% Strain-Peened	120-140	51
Conventionally Shot-Peened and Stress-Relieved	120-140	48

Fig. 7 — Separate effects of surface roughness and residual compressive stress on fatigue properties

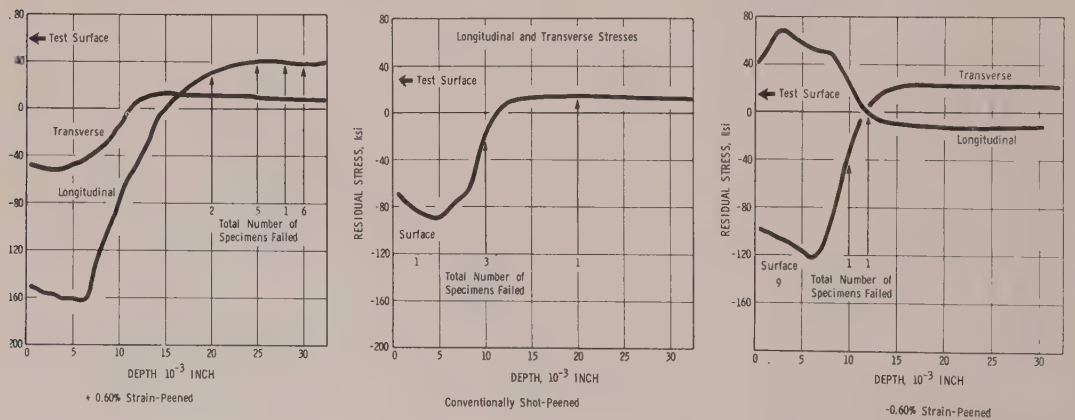


Fig. 8 — Location of fatigue failure origins and residual stress distributions

peened, stress-relieved" group has a more favorable residual stress near the surface, which may account for this observation.

While the linear relationship of the peak residual stress in the affected surface layer and endurance limit shown by Fig. 6 looks impressive, its real meaning is somewhat clouded by the fact that specimens which were tensile strain-peened had their origins of failure below the surface where the stress state was quite different. (See Fig. 8, for instance.) The same applies to specimens which were shotpeened under compressive strain, which actually showed evidence of failure near the surface. Furthermore, extension of the curve to the point of zero endurance limit indicates a mean stress of about 110 ksi. Since the yield strength is nearly twice that much, this does not seem to be realistic. Because of these points, a more thorough analysis was made of the residual stresses in depth in both the longitudinal and the tranverse direction of three types of specimens: (1) peened under 0.60% tension, (2) conventionally shotpeened, and (3) peened under 0.60% compression. Results are shown in Fig. 8, which also gives failure locations. Incidentally, measurements of residual stresses were made both before fatigue testing and after fatigue testing at loads corresponding with endurance limit. These revealed little or no change occurring in the macro-residual-stress distribution because of the fatigue cycling.

Correlation with Failure Criteria — From biaxial residual stress data and the endurance limits observed, it was possible to examine the results in relation to various failure criteria. Consideration was given to the maximum stress theory, the distortion energy theory, and the maximum shear stress theory.

Residual stresses and load stresses were added algebraically (superpositioned). Longitudinal stresses and transverse stresses were considered separately to determine the combined stress state under conditions of endurance limit loading for each of the three groups of specimens (+0.60% strain-peened, conventionally shotpeened, and −0.60% strain-peened). The results are plotted in Figs. 9–11. The

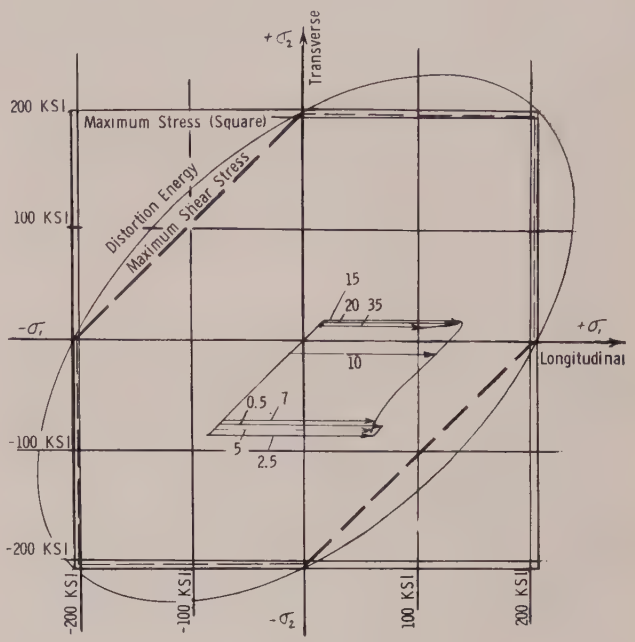


Fig. 10 — Failure criteria and biaxial stresses resulting from superpositioning of residual and endurance limit stresses for conventionally shotpeened group. Stresses are shown for various depths below surface (distance below surface in 1/1000 in. is indicated by numerals on stress envelope)

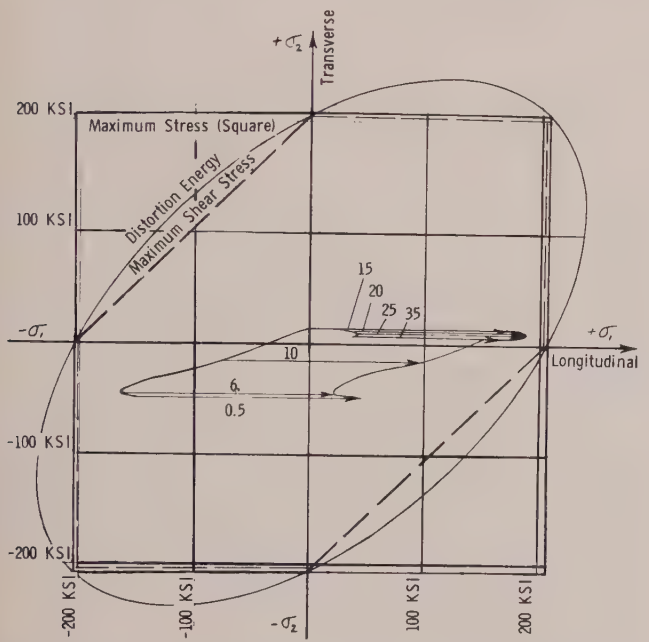


Fig. 9 — Failure criteria and biaxial stresses resulting from superpositioning of residual and endurance limit stresses for +0.60% strain-peened group. Stresses are shown for various depths below surface (distance below surface in 1/1000 in. is indicated by numerals on stress envelope)

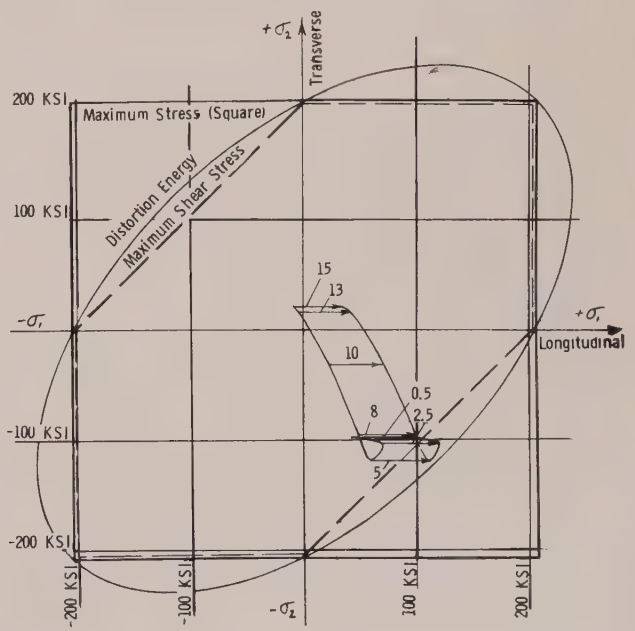


Fig. 11 — Failure criteria and biaxial stresses resulting from superpositioning of residual and endurance limit stresses for −0.60% strain-peened group. Stresses are shown for various depths below surface (distance below surface in 1/1000 in. is indicated by numerals on stress envelope)

three groups are combined in Fig. 12 for purposes of comparison. In each of the individual diagrams the biaxial stress state under conditions of no-load and full-load are shown for near the surface and for various points beneath the surface. The length of the horizontal element in each of these envelopes represents the stress range at the endurance limit. (In the experimental method of fatigue testing used, the transverse stress due to applied load is zero). The loci of the minimum and maximum stresses for each case show relative nearness to the critical stress states defined by each of the three theories.

For purposes of reference the yield point ellipse based upon the distortion energy criterion is shown in true scale. A geometrically similar ellipse should be considered for fatigue, but its scale value is not known. The maximum shear stress diagram and the maximum stress diagram are also shown, but neither of these is in the true scale since the true values are not known.

Considering the maximum stress theory, only stress in a longitudinal direction need be considered. Here, as can be seen from Fig. 12, the tensile-peened group should be the weakest and the compressive-peened group the strongest, yet it is obvious that the opposite is true.

According to the maximum shear stress theory the compressive-peened group should be, and is, the weakest; but the conventionally shotpeened group should be the strongest, and it is not. Its strength is greatly exceeded by the tensile-peened group.

For the distortion energy criteria the predicted order of strength would be: compressive-peened weakest, tensile-peened next, and conventionally peened strongest.

Consideration must be given, however, to possible variations in strength through the cold-worked layer. Predictions would be hazardous, but to satisfy both the distortion energy theory and the maximum shear stress theory the material would have to be weaker at an intermediate depth than either the region near the surface (0–0.005 in. depth) or the region immediately below the cold-worked layer (0.025–0.030 in.). This seems highly improbable.

To satisfy the maximum tensile stress theory the strength of the material below the cold-worked layer would have to be stronger than that within the cold-worked layer. This likewise seems improbable.

Considering all three theories, it appears that either the distortion energy theory or the maximum shear stress theory would satisfy the results better than the maximum stress theory. Both the tensile-peened group and the compressive-peened group are satisfied by these theories. Only the conventionally shotpeened group is not. It may be that better correlation would be obtained if the stress perpendicular to the surface were considered. Furthermore, the maximum stress theory predicts results opposite to those observed.

It is also interesting to note that the failure origins are located subsurface for the tensile strain-peened group, as would be predicted by all three theories, and this might be expected if the surface is "protected" by a compressed layer. All compressive strain-peened specimens had their origins at or very near the surface, as would be expected from the failure theories. The conventionally shotpeened group had origins at various locations, as might be predicted by either the maximum shear stress theory or the distortion energy theory.

It is to be noted, however, that the closer the mean stress is to the critical stress for each of the failure criteria the narrower is the permissible alternating stress. For example, the compressive-peened group has a low alternating stress, and its mean stress is very close to the yield ellipse. The conventionally peened and the tensile-peened follow in that order if conditions of stress near the surface are considered. This is not clearly true beneath the surface, but variations in strength may be responsible for this divergence.

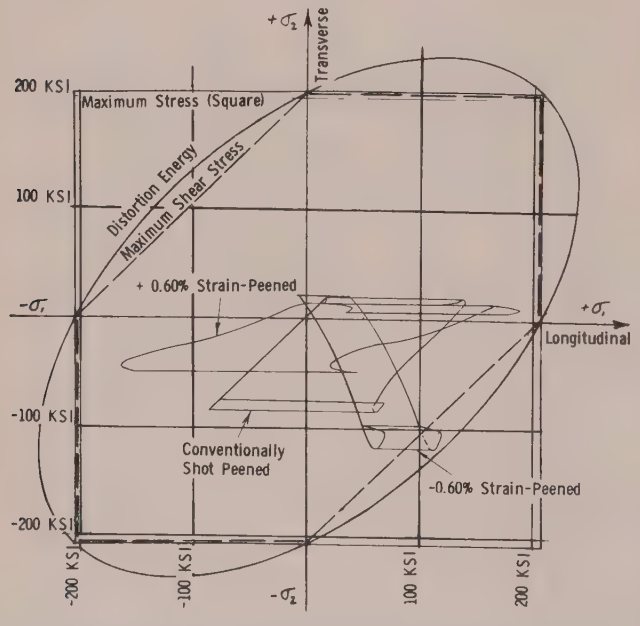


Fig. 12 — Combined biaxial stress diagrams of three groups: +0.60% strain-peened, conventionally shotpeened, and −0.60% strain-peened, taken from Figs. 9-11

Conclusions

For the material, specimen, and testing described in this experiment, the following conclusions are drawn:

1. Good correlation was found between endurance limits and the residual stress state near the surface independent of any strain-hardening influence. This supports the concept of superpositioning of residual stresses and load stresses. True quantitative interpretation was not possible with the limited data, but none of the results contradict the idea of considering residual stresses as being additive to mean load stresses.

2. It appears from this investigation that shotpeening steel of this hardness is indeed beneficial primarily because of the introduction of the macro-residual-compressive-stress. There is no gain attributable to "strain-hardening" for this material. There is some indication that the increased surface roughness caused by shotpeening is harmful to fatigue strength.

3. Fair correlation was found between endurance limits and the distortion energy criterion and the maximum shear stress criterion of failure using the principle of superpositioning of residual stress and load stresses and failure locations.

Torsionally Resilient Drive Lines

P. J. Mazziotti
Dana Corp.

THE RECENT increased use of midship-mounted transmissions in large equipment has emphasized the need for a more thorough engineering approach to the drive line torsional resonance problem. It was found in many installations that the standard, reasonably torsionally stiff drive line connection originally installed produced torsional resonance from engine firing impulses within the operating range. The effect of torsional resonance was always disturbing noise and vibrations and often fatigue failures because of a large peak torque at reasonably high frequency.

In order to reduce vibration transfer, it became necessary to provide a more resilient connection between engine and transmission. This installation could actually allow for operation well above resonant frequency. The post-resonant condition results in the minimum disturbance and reduces peak loads.

Torsionally resilient drive line constructions are not new. Many different forms have been used over the years. Not all of these units have been properly installed; they were, in some cases, the result of trial-and-error installation. The calculations necessary to determine the proper torsional rates to suit units are simple. The analysis developed in this paper is intended to be a practical approach to the problem.

The calculations are pointed towards the problem of midship-mounted transmission installations; however, the analysis used can be applied to a wide range of installations. These are the same principles that have been used in engine dynamometer installations, compressor drives, and the like where nonuniform rotating motion occurs.

Simplified Installation

The installation of a drive line between two rotating inertias can be simplified for purposes of analysis as shown in Fig. 1. In this sketch the two rotating inertias I_1 and I_2 are connected with a spring of rate k. The torsional spring can be of several forms. The illustration shows a simplified solid shaft connection and a torsional rubber spring.

For purposes of analysis, the theoretical torsional vibrating system of a single rotating inertia I connected through a spring of rate k to a rigid foundation is also shown. The simplified system is used for a basis of comparison with the two rotating inertia systems actually under consideration.

In actual installations the two rotating inertias can be from several possible units:

1. Engine, motor, or other prime mover.
2. Transmission, torque converter, fluid clutch, and the like.
3. Axle gearing, differentials, and the like.
4. Pumps, compressors, fan, blowers.
5. Generators, alternators, and so forth.

In some cases, the effective rotating inertia may

THE INCREASED use of midship-mounted transmissions in large equipment has emphasized the need for a torsionally resilient connection from the engine to reduce vibration transfer. To increase the torsional flexibility needed in these systems, the spring rate of the system must be reduced by such constructions as a flexible coupling, a spring-loaded damper, or a rubber torsional spring.

This paper discusses these systems, emphasizing rubber springs. Some advantages of such a drive are: it provides an amplitude limitation with impact loads and a cushion to reduce noise and prevent clattering and contacts noises on parts with backlash, it smooths out transition periods to reduce loads on bearings and gears, its clamping characteristics can be adjusted by various rubbers, and its rubber cushion provides a degree axial flexibility.*

* Paper presented at Earthmoving Conference, SAE Central Illinois Section, Peoria, April 14, 1959.

vary somewhat with speed and load. In the case of a hydraulic torque converter pump, the effective inertia increases somewhat as torque load increases which can affect the resonant frequency. The analysis can be modified to take advantage of this effect, if desired.

The two rotating inertias of the end units connected by a torsional spring will have a natural torsional frequency. When a source of nonuniform motion exists in the system, torsional resonance can be established. The torsional impulses from nonuniform motion will be transferred through the rotating system and be amplified or subdued, depending upon frequency of disturbance and the natural frequency of the system. Fig. 2 shows how the disturbance can be magnified or subdued at various frequencies. From this curve the following conclusions can be made:

1. It is assumed that the primary disturbing force has the magnitude of one. In actual applications, the value of the disturbing force will, of course, change with rotating frequency.

2. The peak effective transferred force occurs at the resonant frequency. The actual amplitude at this point depends upon the amount of damping present. A change in damping will effect a large change at resonance. The effect will diminish as frequencies depart from resonance.

3. In order to transmit less disturbing force than is fed in to the system, the rotating speed must be over 1.5 of resonance.

4. The disturbing force reduces very slowly above two times resonance.

The amplification effect upon the disturbing force is shown as a steady state condition. When the resonant frequency is passed through quickly, the limited number of cycles may not be able to achieve the maximum steady state condition shown. This often allows passing through resonance frequency with much less magnification than would be present in steady operation.

Possible Sources of Torsional Excitation

There are several possible sources of torsional excitation which can produce the forced frequencies and resonant condition.

1. Nonuniform motion resulting from firing and compression cycles in a reciprocating engine. This is the primary case under consideration in the engine-to-transmission installation. It is expected that their firing impulses will cause a torsional acceleration. The magnitude of acceleration will depend upon the inertia of the flywheel and other rotating components relative to the magnitude of the firing impulses. The possibility exists that a single cylinder may not fire and cause a lower frequency excitation than that resulting from uniform firing. Fig. 3 shows the possible motion resulting from uniform firing of a 6-cyl, 4-stroke engine and the possible lower frequency with one cylinder misfiring. Uniform firing produces three impulses per revolution. When one cylinder misfires, a fundamental frequency exists of one-half cycle per revolution. There exists the possibility of all other multiples of these frequencies.

2. A similar excitation can come from other units because of their torsional characteristics:

(a) Impulses from other reciprocating machinery such as compressors.

(b) Variable load characteristics of a periodic nature.

(c) Nonuniform motion geometry such as the two per revolution nonuniformly existing from a universal joint operating at an angle.

(d) Nonuniform contact of gears, cams, linkages, or other motion transferring devices.

3. The amplitude of exciting vibration can be changed within the end units. The addition of a flywheel to a nonuniform rotating unit will have the effect of reducing the amplitude of disturbances. The response characteristics of the entire driving system also effects the amplitude of excitation present.

It is apparent that there are many possible sources of nonuniform rotating motion which can produce a disturbing force on a rotating system.

Torsional Elastic Properties

The simple installation of a drive line between two rotating inertias can be assumed to be a torsional spring resulting from the elastic properties of the steel connecting shaft. The torsional rate of this steel shaft can be determined from the following equations:

$$k = \frac{\Gamma}{\theta} = \frac{E_s J}{L} \qquad (1)$$

where:

k = Spring rate, in.-lb per radian
θ = Angle of twist in radians
Γ = Torque, in.-lb
E_s = Shear modulus lb per in.^{2} (12 million for steel)
J = Polar section moment of inertia
L = Effective length, in.

The stress in the shaft can be determined from:

$$S_s = \frac{\Gamma D_o}{2J} \qquad (2)$$

where:

S_s = Maximum shear stress
D_o = Outer diameter of shaft

From these equations, it is apparent that spring rate can be decreased (flexibility increased) by the following changes:

1. Raise working S_s. In practice the shear stress must be kept reasonably low in order to absorb peak torques resulting from rapid changes in rotating speeds.

2. Increase length. The actual length is usually limited by installation requirements. In all too many cases, the allowable length is a minimum necessary to put in the connecting unit. If the effective length could be increased, it would still be limited by other considerations such as critical speed.

3. E_s is a constant value for steel; however, other materials have a lower value which could be an advantage. In general, most materials with a lower modulus will have a correspondingly lower working stress which may not result in any overall gain.

4. Reduce the polar section moment of inertia to a minimum. This will lead to the conclusion that a solid shaft produces the maximum flexibility at a given stress level.

The torsional flexibility can be increased by the addition of auxiliary resilient members of several possible forms. These auxiliary resilient members

can be in the form of a flexible coupling, spring loaded damper, rubber torsional spring, and many possible combinations thereof. These possibilities will be discussed in more detail later.

Resonant frequency can be reduced by increasing the rotating inertias on one or both ends of the torsional spring. The effectiveness of this is limited and it can be done only at a considerable weight and cost penalty.

Theory and Analysis of Natural Frequency

The mass-elastic system consisting of two rotating inertias connected with a torsional spring can be analyzed based on the following:

1. Both inertias vibrating at the same frequency.
2. Vibrations are out of phase. This results in a nodal point between them.

$$\frac{1}{k} = \frac{1}{k_1} + \frac{1}{k_2} \tag{3}$$

where:

k = Spring rate in.-lb per radian
k_1 = Spring rate I_1 to node
k_2 = Spring rate I_2 to node

The general equation for the natural frequency of a mass of inertia I on a torsional spring rate k attached to a rigid support is:

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{I}} \tag{4}$$

where:

f = Natural frequency, cps
I = Mass polar moment of inertia, lb-in. sec^2

Since both inertias are vibrating at the same frequency

$$\frac{k_1}{I_1} = \frac{k_2}{I_2} \tag{5}$$

Substituting this relationship into the natural frequency equation gives the natural frequency for the two mass systems as:

$$f = \frac{1}{2\pi}\sqrt{\frac{k\,(I_1 + I_2)}{I_1\ \ I_2}} \tag{6}$$

To compare the natural frequency of the one- and two-mass systems, a set of constant values can be assumed and the relationships determined. Assume:

$$\frac{\sqrt{k}}{2\pi} 1 \qquad\qquad I = I_1 = 1$$

With these values

$f = 1$ for the one-mass system (7)

$$f = \sqrt{\frac{1 + I_2}{I_2}} \text{ for the two-mass system} \tag{8}$$

Fig. 4 shows how the frequency of the two-mass system compares with the frequency of the single-mass system with various values of $\frac{I_2}{I_1}$. Some conclusions can be made on the effects of changes in the two-mass system as follows:

1. Increasing either I_1 or I_2 will lower the natural frequency.
2. If one inertia is large compared to the other one, increasing the large inertia has a comparatively small effect on the natural frequency.
3. If one inertia is large compared to the other, increasing the smaller inertia has a comparatively large effect.
4. Overall reduction in frequency is small if a change is made in either inertia where they are of approximately the same size.

There are practical limitations in the ability to change torsional resonant frequency of a two-mass system in changing the size of the rotating inertias. This approach may be practical where one inertia is very small compared to the other. Where more equally sized inertias are involved, a change in resonant frequency can be accomplished primarily by change of spring rate.

Effects on Engine-to-Transmission Installations

In a drive line from an engine to a transmission, a torsional resonance condition can be established

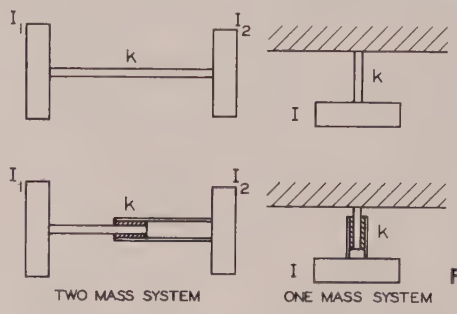


Fig. 1 — Mass elastic system

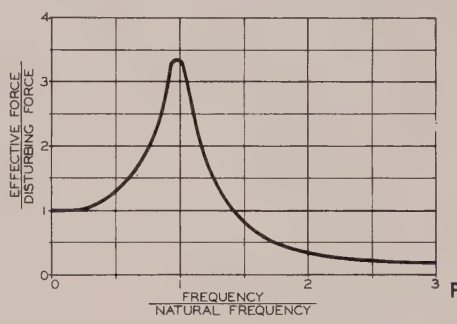


Fig. 2 — Magnification of disturbing force

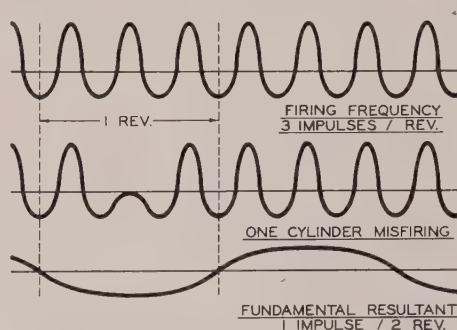


Fig. 3 — Firing impulses from 6-cyl 4-stroke engine

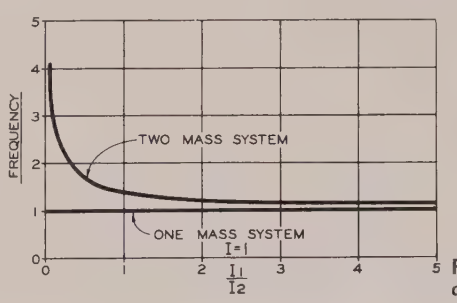


Fig. 4 — Natural frequency versus inertia

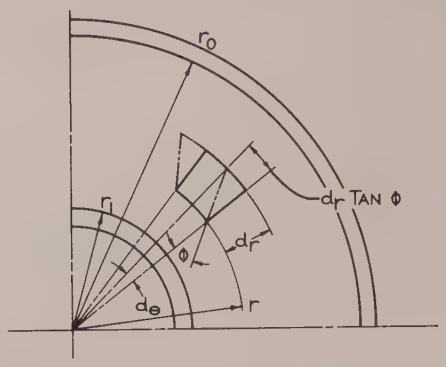


Fig. 5 — Relationships under load in rubber torsional spring

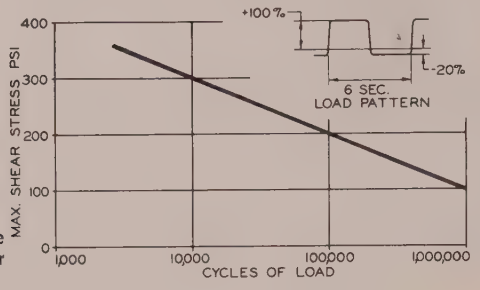


Fig. 6 — Load-life curve for rubber torsion springs

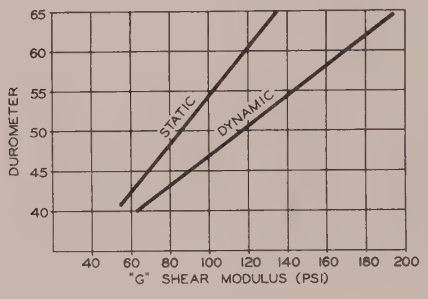


Fig. 7 — Rubber durometer versus shear modulus

where engine impulses coincide with natural torsional frequency of the installation. The actual peak oscillating torque will depend on the characteristics of the system and damping present in the entire unit. A resonant condition can produce objectionable disturbances as follows:

1. The high oscillating torque values can result in fatigue failures in rotating members.
2. Variable reactions on supporting members can be a source of objectionable noise and vibration.
3. Damage to gears, bearings, and other components because of nonuniform loading.

Fig. 2 shows that at all frequency under 1½ times the natural frequency, some magnication of disturbance will occur. In order to transmit a disturbance appreciably less than the disturbing force, the frequency should be at least two times the natural frequency. Above this, the reduction in transmitted disturbance is slow.

Torsionally Resilient Units

To increase the torsional flexibility (reduce spring rate) in a drive line, several constructions can be considered.

Flexible Couplings — Several types of flexible couplings are available but have many limitations in use. They are usually limited to reasonably small

Table 1 — Drive Line Installation

Engine:	6 cyl, 4-stroke Diesel 850 ft-lb maximum net torque 2500 rpm maximum operating speed 500 rpm approximate idling speed $I_e = 40.5$ lb in sec^2
Transmission:	Hydraulic torque converter and planetary gear unit midship mounted $I_t = 22.2$ lb in sec^2 effective on the transmission input shaft
Installation Limitations:	Length available 16½ in. flange face to flange face Diameter—approximately 9-in. swing diameter maximum for clearance

total torsional wind up (approximately 7 deg maximum at full capacity). They are usually quite large in size and weight for their capacity.

There is oftentimes a dynamic balance problem associated with their use because of the flexibility and instability inherent in resilient members. Wherever a reasonably large amount of continuous misalignment is present, the resilient member may not be capable of sustained high-speed operation and the required continuous deflections. The life of most of these units goes down rapidly as the operating misalignment increases.

There may be many installations that a flexible coupling provides a reasonable amount of torsional resilience and fits the installation requirements. It is beyond the intent of this paper to discuss the pros and cons of many different designs and constructions which can be considered.

Spring-Loaded Damper — This construction provides a satisfactory way of increasing torsional flexibility provided it can be fitted into the installation. There is some problem in obtaining large amplitudes with low rates because of design limitations. It is, of course, possible to use elastomer springs in place of mechanical springs; however, this does little to simplify the amplitude problem. The mechanical spring type oftentimes provides an additional noise and clatter problem because of mechanical contacts under oscillating loads. The spring-loaded damper resilient device is a reasonable unit for consideration in increasing torsional flexibility. They are reasonably difficult to fit between two universal joints or couplings because of their size and shape.

Rubber Torsional Springs — The size and capacity relationship of these units make them a very acceptable way of obtaining increased torsional flexibility between universal joint centers. With reasonable proportions the diameter of the unit can fit within universal joint diameter. They have many desirable features which lend to their desirability as a resilient unit:

1. Can be made to oscillate through large amplitudes.
2. Between universal joint centers the rubber torsional spring must carry only transmitted torque and remaining oscillations without having to flex through high frequency misalignment.
3. Can fit between universal joint centers away from the highest temperatures and installation problems at each end.
4. With reasonable proportions they can be lightly stressed for long life.

The characteristics of a rubber torsional spring of suitable proportions for fitting between universal joint centers will be considered here in some detail.

Fig. 5 shows the relationships which exist in a rubber spring under a torsional load. From these relationships the equations used in calculating the characteristics are as follows:

$$S_{mox} = \frac{\Gamma}{2\pi r_1^2 L} \tag{9}$$

where:

S = Shear stress psi at a radius r
r_i = Inside radius
r_o = Outside radius
L = Length of rubber

From Fig. 5:

$$r d\theta = dr \tan \phi \tag{10}$$

$$\phi = \frac{S}{G} = \frac{\Gamma}{2\pi r^2 LG} \tag{11}$$

where: G = Shear modulus of rubber

$$d\theta = \frac{1}{r} \mathrm{Tan} \left[\frac{\Gamma}{2\pi r^2 LG} \right] dr \tag{12}$$

A tangent function can be expressed as a series solution as follows:

$$\mathrm{Tan}\, X = \left[X + \frac{X^3}{3} + \frac{2X^5}{15} + + \right] \tag{13}$$

Since this series converges rapidly a reasonable approximation can be made by using the first term only. Using this simplified solution:

$$d\theta = \frac{\Gamma dr}{2\pi r^3 LG} \tag{14}$$

$$\theta = \frac{\Gamma}{2\pi LG} \left[-\frac{1}{2r^2} \right]_{r_i}^{r_o} \tag{15}$$

$$\theta = \frac{\Gamma}{2\pi LG} \left[\frac{1}{2r_i^2} - \frac{1}{2r_o^2} \right] \tag{16}$$

$$k = \frac{\Gamma}{\theta} = \frac{4\pi LG r_i^2 r_o^2}{r_o^2 - r_i^2} \tag{17}$$

In order to make use of these equations, some working values must be established for stress, modulus, and design proportions. A practical value for maximum shear stress in the rubber torsional spring can be established by analysing the static and fatigue capacity of the rubber unit. In connecting the engine to the transmission where operation is primarily over two times torsional resonance frequency, a very small part of the torsional disturbance is actually transferred through the connecting drive line. It can be reasonably assumed that if continuous operation at torsional resonance is avoided that fatigue life can be based primarily upon maximum engine torque.

An approximate load-life curve for a rubber torsional spring is shown in Fig. 6. The actual load-life relationship in this type unit varies considerably depending upon many variables as follows:

1. Frequency of Load

Because of internal hysteresis the temperature will increase if fatigue cycles are at a reasonably high frequency. It has been found that tests can be slowed down to the order of one cycle of loading per 6 sec to keep maximum temperature low.

2. Load Pattern

The shape of the load-time curve has an effect on the overall life of the unit. The results shown in Fig. 6 are based upon an approximate square wave loading of plus 100% of test load to minus 20% of test load as shown on the illustration.

3. Rubber Properties

The strength of the rubber material, its bond strength at the point of maximum stress, the amount of internal damping, all have an effect on its fatigue life. Material tested to obtain the load-life curve shown in Fig. 6 is in the range of 45–55 durometer natural rubber with comparatively low internal hysteresis.

4. Test Procedure

The fatigue test data obtained for Fig. 6 was obtained by fatiguing at low frequencies by successively applying torque load through an air motor drive. This type of unit does not wind up a spring to a fixed amplitude, it has the property of stopping at a given load at any amplitude where it occurs. It is assumed that a failure occurs when the original amplitude has doubled.

Test results show that bond strength at the ID is the critical working stress limitation. To obtain long fatigue life at maximum engine torque, it would appear that S_s maximum should be designed for about 100 psi. This might be increased in some cases where maximum torque can occur only a few times and impulses are kept very low.

A working value for G (shear modulus of the rubber) can also vary considerably with natural physical properties of the elastomer. The curve shown on Fig. 7 shows approximate relationship between durometer and static and dynamic shear modulus. These values are reasonably accurate and offer a sound basis for design analysis. The dynamic modulus is of primary concern in establishing natural frequency values. From this curve it can be seen that the shear modulus will be in the order of 100 psi in the 45–55 durometer range.

These equations and curves provide enough information to analyze the effects of a rubber torsional spring in a transmission to engine installation. A comparison can be made between the effects of using a short steel shaft unit which produces objectionable torsional resonance frequency and the improvements possible with the torsionally resilient rubber spring.

Typical Problem

An analysis is to be made of an engine-to-transmission drive line for a large size vehicle. Calculations are to be based primarily on avoiding torsional resonance as much as possible. The data for the installation is shown in Table 1.

A short length propeller shaft assembly can be used as a drive line between engine and transmission. This would represent the simplest and lowest cost installation. A universal joint of approximate 1700 Series capacity would be suitable for this installation with its 8 in. maximum swing diameter and suitability for the 16½ in. installed length. If it is assumed that a propeller shaft of this type could be modified to use a minimum practical size solid shaft connected between the end units, the following proportions can be considered:

1. Effective length of solid shaft 13½ in. This allows 3 in. for ineffective length due to the comparatively stiff members in the universal joint.

2. Solid shaft shear stress to be maintained at 10,000 psi to allow suitable factor for peak loads

(estimated at 5 to 1) and obtain satisfactory fatigue life.

$$D^3 = \frac{16\Gamma}{\pi S_s} \tag{18}$$

where: $D = 1.74$ in. diameter based on values above

3. Torsional rate can be determined by:

$$k = \frac{\Gamma}{\theta} = \frac{E_s J}{L} \tag{19}$$

where: $k = 792{,}000$ in.-lb per radian

4. The natural frequency of the system can be determined by:

$$f = \frac{1}{2\pi}\sqrt{\frac{k(I_e + I_t)}{I_e I_t}} \tag{20}$$

where: f = 37.4 cps or 2250 cpm

With three cycles per revolution from firing frequency, torsional resonance occurs at 750 rpm. If two times resonance is considered to be a desirable minimum condition, the system can be considered to be unsuitable for operation under 1500 rpm.

With one cylinder misfiring, resonance will occur at 4500 rpm. This is above the operating range and would result in a reasonably large amplitude transfer through the entire range. These resonance values cannot be considered satisfactory for this installation.

A completely ideal installation would have a torsional resonance frequency well under half of engine idling speed with one cycle per two revolutions excitation. In order to achieve resonance at 150 cpm (2.5 cps), the spring rate of the connecting shaft would have to be 3540 in.-lb per radian. This is an extremely low spring rate and would require a unit which would wind up 2.9 radians (166 deg) at full engine torque. This amplitude is extremely high and probably cannot be obtained.

A practical set of proportions for a rubber torsional spring desirable within these limitations would be as follows:

Inside diameter = 4 in. ($r_i = 2$ in.)
Outside diameter = 8 in. ($r_o = 4$ in.)
$S_s = 100$ psi maximum at 850 ft-lb torque

$$L = \frac{\Gamma}{2\pi r_i^2 S_s}$$

$L = 4.07$ in. (use 4½ in.)
G (dynamic) = 90 psi based upon 45 durometer rubber (Fig. 7)

$$k = \frac{4\pi L G r_i^2 r_o^2}{r_i^2 - r_o^2}$$

$k = 27{,}100$ in.-lb per radian

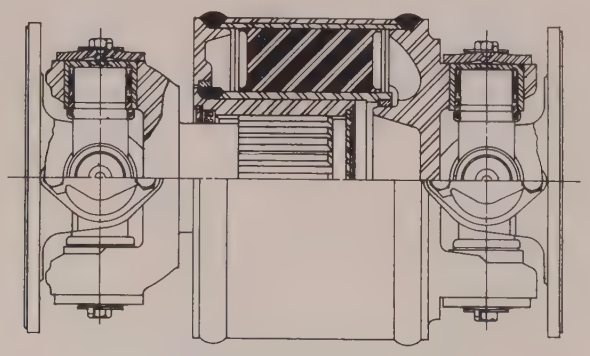


Fig. 8 — 1700 universal joints with rubber torsional spring

The natural torsional frequency with this spring rate will be

$f = 6.9$ cps or 416 cpm

At firing frequency, torsional resonance would occur at 139 rpm which is well below idling speed. Under extreme condition of one cylinder misfiring, resonance frequency would occur at 832 rpm which is in the operating speed; however, probably would not cause a failure since it could be passed through comparatively fast.

Fig. 8 shows proportions of a 1700 Series propeller shaft that would fit these requirements. Standard universal joints are modified to fit large torsional rubber spring. The short length of the rubber unit makes it comparatively flexible in resisting sideways forces. The tilting of the rubber unit can seriously increase unbalance and produce rough operation. To protect against this, auxiliary supports around the rubber unit provide additional support. The finished package produces a clean, simple, practical unit to accomplish desired torsional flexibility.

Possibilities of Expanded Use

The addition of torsional flexibility in a drive line is most urgently needed where a torsional resonant condition is to be avoided. The addition of flexibility does offer additional advantages and desirable operating characteristics outside of the resonance considerations. These additional advantages are as follows:

1. The torsional flexibility provides an amplitude limitation on impact loads such as rough shifts and sudden changes in torque which cushion, modify, and smooth out the transition period.

2. The torsional flexibility produces cushion which tends to prevent clattering, rattling, and contact noises on parts with backlash.

3. The smoothing out of transition periods, impacts, and clatters will reduce peak loads on bearings, gear teeth, splines, and other components which will result in increased life. Increased universal joint life is one of the gains that can be expected when the torsional flexibility is added.

4. The rubber unit itself acts as an insulator in reducing noise transfer.

5. Damping and internal energy dissipation characteristics can be adjusted by various rubber characteristics.

6. The rubber cushion unit also provides a degree of axial flexibility which can cushion forces resulting from changes in length.

There are, of course, a number of disadvantages to a rubber torsional spring unit in a drive line:

1. Increased cost.

2. More difficult to balance especially with short length rubber unit.

3. Increased weight.

4. Danger of improper installation where torsional resonance is brought down within the operating range. It could be well above the operating range without the resilient member.

There are many types of installations where torsional resonance presents a serious problem. The proper use of torsionally resilient units to place this resonance frequency outside of operating range is promising. It is to be expected that the use of properly installed resilient members will increase as advantages in their use become more apparent.

prediction in new

Metal Joining Processes

NEW WELDING processes are dropping costs while providing improvements in weld quality. This paper describes some of the more promising new developments in pressure and fusion welding and brazing.

Included in the discussion are ultrasonic, high frequency resistance, foil seam, magnetic force, percussion, friction, and thermopressure welding and diffusion bonding. The description of adhesive bonding includes the development of glass or ceramic materials as structural adhesives.*

John J. Chyle

A. O. Smith Corp.

THE FIELD of metallurgy has seen many new developments in recent years. New metals and alloys which were nonexistent a decade ago are now being produced in quantity. While 50 years ago only a few welding proceses were known and in actual production, today we have over 60 distinct processes. These have been developed in a few highly specialized fields; a number have been created because of developments in electronics and electrical engineering.

Joining of metals may be classified into the following three categories:

1. Welding.
2. Adhesive bonding.
3. Mechanical fastening.

The discussion will be limited to the most recent developments in welding, describing some of the important advances in adhesive bonding which have been reported in literature.

Welding can be defined technically as a process in which localized coalescence of metal is accomplished, wherein the coalescence may be produced entirely by pressure with or without the accompaniment of heat, or where coalesence is produced by heating alone with or without the application of pressure, and with or without the use of filler metal.

Table 1 lists the classifications of weld proceses. In each of the subdivisions listed, there are additional subclassifications of welding processes, depending on the source of heat and the type of shielding gases. The total available welding processes exceed 65 separate and distinct processes. The selection of any process will be governed by such factors as metals to be joined, design of weldment, and cost of welding.

Pressure Welding

Pressure welding may be defined as a group of related welding processes characterized by the use of force as being of primary importance in obtaining the coalescence of metal, which usually is performed over the entire area of the abutting surfaces. This

* Paper presented at SAE National Aeronautic Meeting, New York, April 1, 1959.

Table 1 — Classification of Welding Processes

I. Pressure Welding Processes
1. Forge welding
2. Cold welding
3. Resistance welding
4. Pressure thermit welding
5. Solid phase welding
6. Pressure arc welding
7. Explosive force welding
8. Ultrasonic welding
9. High frequency resistance welding
10. Foil seam welding
11. Magnetic force welding
12. Percussion welding
13. Friction welding
14. Thermo pressure welding
15. Diffusion bonding

II. Fusion Welding Processes
1. Arc welding—carbon electrode
2. Gas welding
3. Arc welding—metal electrode
4. Braze welding
5. Flow welding
6. Arc welding—tungsten electrode
7. Atomic hydrogen welding
8. Submerged arc welding
9. Induction welding (nonpressure)
10. Electron beam welding
11. Electro-slag welding
12. Arc plasma welding

III. Brazing
1. Torch brazing
2. Furnace brazing
3. Dip brazing
4. Flow brazing
5. Twin carbon arc brazing
6. Resistance brazing
7. Block brazing
8. Induction brazing
9. Vacuum brazing
10. Radiant heat brazing

Table 2 — Typical Strengths of Ultrasonic Single-Spot Welds in Several Materials[a]

Material	Type	Sheet Gage in.	Tensile-Shear Strength[b] lb
Aluminum	2024-T3 Bare	0.063	1,290 ± 60
Aluminum	2014 T6	0.063	1,310 ± 120
Aluminum	2020-T6	0.040	1,240 ± 50
Aluminum	7075-T6 Alclad	0.050	1,540 ± 90
Mild Steel	1020	0.020	500 ± 20
Stainless Steel	Type 301	0.025	1,040 ± 70
Stainless Steel	17-7 PH	0.020	668 ± 110
Molybdenum	Powder metallurgy	0.025	430 ± 50
Molybdenum	Arc cast	0.025	360 ± 40
Inconel		0.010	170 ± 70
Inconel		0.020	220 ± 105
Titanium	A-110AT (5 Al-2.5 Sn)	0.028	1,950 ± 120
Titanium	C-120AV (6 Al-4V)	0.025	1,230 ± 60
Titanium	C-110M (8 Mn)	0.032	1,730 ± 200
Zircaloy-2		0.020	650 ± 20

[a] "Ultrasonic Welding Techniques," by J. Byron Jones, and W. C. Potthoff. American Society of Tool Engineers, Vol. 58.
[b] Average value with 90% confidence intervals.

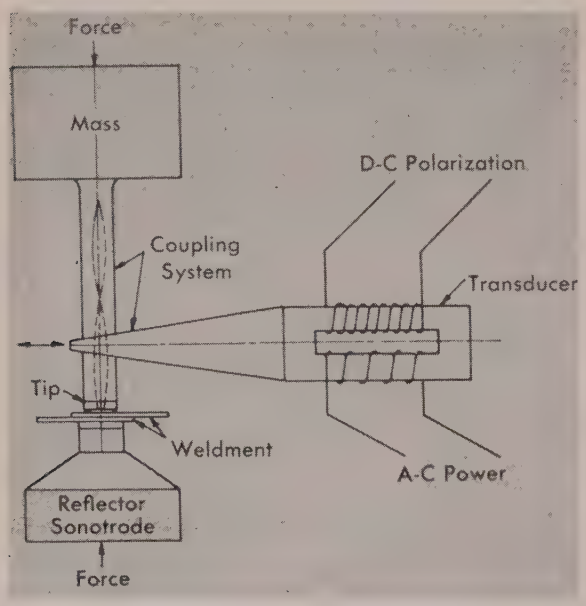


Fig. 1 — Diagrammatic sketch of component parts for ultrasonic welding process (from Metal Progress, April, 1958, J. Byron Jones)

coalescence may be produced either with or without the application of heat.

Where heat is applied, these processes may be further subdivided as to the source and nature of the heat—whether it is from an electrical source, a chemical reaction, or the combustion of solid fuels or gases. In some of the pressure welding processes there is some overlapping where fusion of the metals is accomplished. However, these processes are classified under pressure welding, since pressure is necessary to produce a satisfactory weld.

Ultrasonic Welding

The ultrasonic welding process has received considerable attention during the last few years, and appears to have promising applications for future use. In this process the energy to effect a weld is derived from the use of special equipment, using elastic vibratory energy in the joining of metals. Both spot and seam welding of metal sheets can be accomplished using commercially developed equipment.

Fig. 1 is a diagrammatic sketch of the component parts, and consists of a transducer and a coupling system which is incorporated to a vibratory element or an electrode. The weldment is inserted between the electrode tips as shown.

Another form of ultrasonic welding is shown in the continuous seam welder (Fig. 2). With this equipment, ultrasonic energy is applied continuously to produce a pressure tight welded seam. This seam welder has been developed for the welding of thin foil as well as light sheet metal. Both of the rollers shown are powered, rotating at the proper speed to provide the desired traversing speed of the sheets to be welded. In some designs the welding head rotates and traverses along a fixed work piece. In other designs the welding head rotates while remaining in a fixed position and the work piece travels on a moving fixture under the rotating welding head. Metals difficult to weld by conventional methods can often be joined with this process, such as stainless steel, molybdenum, zirconium and its alloys, inconel, tantalum, and titanium. Ultrasonic welding is also used for the joining of not only similar, but also dissimilar metals such as copper to aluminum. Various types of bimetal combinations can be joined, such as:

Brass, copper, steel, and aluminum to zirconium.

Platinum, molybdenum, brass, inconel, and silver manganese to various steels.

Titanium, copper, gold, and platinum alloys to nickel.

Copper or silver to brass.

Molybdenum or tantalum to aluminum.

Platinum and ruthenium alloys to a silver-gold alloy.

The quality of an ultrasonic weld made between two dissimilar alloys is shown in Fig. 3. This photomicrograph illustrates the junction line between the 0.010 in. thick Type 316 stainless steel and 0.020 in. Zircalloy-2 alloy. Fig. 4 illustrates the use of ultrasonic equipment for fabrication of a shell section.

Substantial progress has been made in the development of the ultrasonic welding process in the past few years. Notable improvement in the strength of spotwelds has been obtained. This has been largely the result of improved design of equipment and the procedure used. At the present time the limitations of this process are for the joining of light gage or thin sheet materials. Foil thicknesses can be joined readily, but there is some limitation as the gage or thickness of the sheets is increased. Recent tests have been made in sheet metal thicknesses of 0.090 in., and spot strengths substantially exceed the Military Specification requirements. Typical strengths of ultrasonic single spotwelds are shown in Table 2. With further developments in the equipment for ultrasonic welding it is expected that the range of thickness will be extended, and that the applications for this process will be expanded.

High Frequency Resistance Welding

This is one of the newer resistance welding processes which depends on high frequency electrical energy to obtain the desired heating characteristics for welding. This process depends on the fact that high frequency current can be brought directly into the object to be heated by contact, as well as by induction. The use of contact electrodes to bring high frequency current into a piece of work results in several peculiar effects. All are due to the fact that the current flows in the low inductance path, but not necessarily in the low resistance path.

In the high frequency resistance heating process, alternating current generated from 100,000 cycles and upwards is connected directly to the work which is to be heated by means of contact electrodes, the current being arranged to flow in almost any desired path which often is not necessarily the low resistance path. At such high frequency, currents flow in the low inductance path as against the low resistance path by various unusual and special arrangements of conductors, according to the product to be welded.

This process was first commercially developed by the New Rochelle Tool Corp. in 1950. Their patented method is based on these high frequency heating principles. The contacts are placed on each side of two abutting surfaces of metal. The edges touch at one end and the rest of the joint is open slightly so as to form a sharp *V*, as shown in Fig. 5. The current flowing from one contact hugs the surface of the other edge of the *V* as far as the root, and completes the circuit by returning along the opposite edge of the *V* with the second contact. This skin effect produces localized high intensity heat-

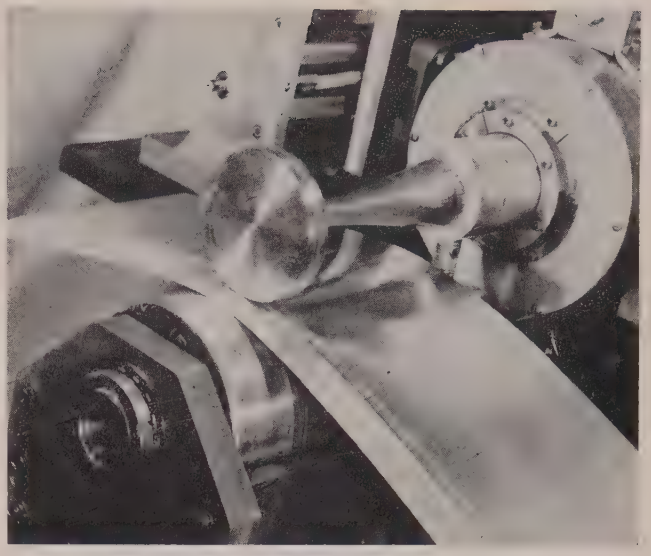


Fig. 2 — Ultrasonic welding equipment for continuous seam welds (courtesy: Aeroprojects, Inc.)

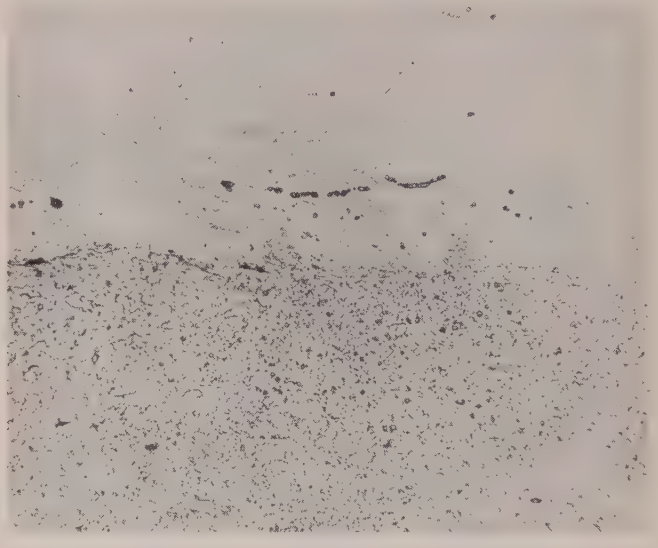


Fig. 3 — Microstructure of ultrasonic weld made between 0.010-in. Type 316 stainless steel and 0.020-in. Zircaloy-2. Original magnified 100 times (courtesy: Aeroprojects, Inc.)

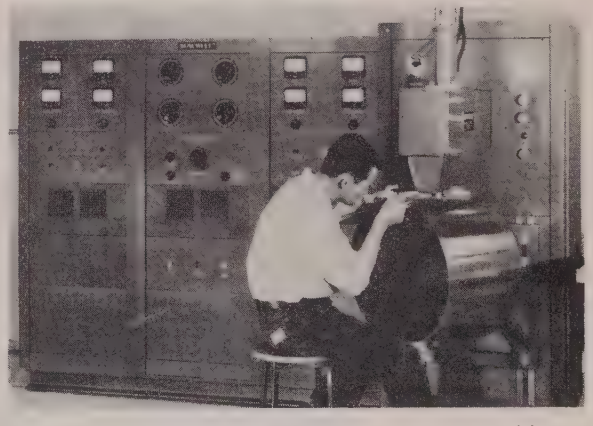


Fig. 4 — Ultrasonic welding equipment as applied for fabrication of shell section (courtesy: Aeroprojects, Inc.)

ing, so that the edges at the root of the *V* reach welding temperatures at the same rate as the strip progresses through the welding machine. Pressure applied at the root of the *V* brings the two edges together in a forge weld. Fig. 6 is a close-up view of the high frequency welding process applied to 6-in. aluminum tubing. Travel speed of the pipe can be up to 200 fpm, depending upon the thickness of the wall. A photomicrograph of a butt weld made in low carbon 0.045-in. wall tubing by the high frequency resistance process is shown in Fig. 7. The weld reinforcements can be removed by hot trimming, depending on the requirements.

This process is applicable to a wide variety of metals, such as aluminum, brass, copper and copper alloys, stainless steels, silicon bronze, Monel, inconel, and incaloy. Some of the highly reactive metals such as titanium, titanium alloys, and zirconium alloys have been successfully welded with this process, using inert gas for shielding. In addition to the welding of similar metals, combinations of dissimilar metals have also been successfully welded, such as copper to aluminum. One of the advantages of this process is the high speed that can be used. Speeds up to 1000 fpm can be obtained in light gage materials of 0.012 in. thickness. The process is not limited to tubing, but can be applied to other shapes such as strips, "T" sections, and the like. Spiral strips of metal can be welded to the surface of tubes, as shown in Fig. 8. Another advantage of this process is the narrow heat-affected zone obtained due to the relatively narrow band of heating. This narrow or localized heating also accounts for the greater efficiency of the electric power used. The narrow weld zone may result in high tensile strengths and ductility. By utilizing high power density, high fre-

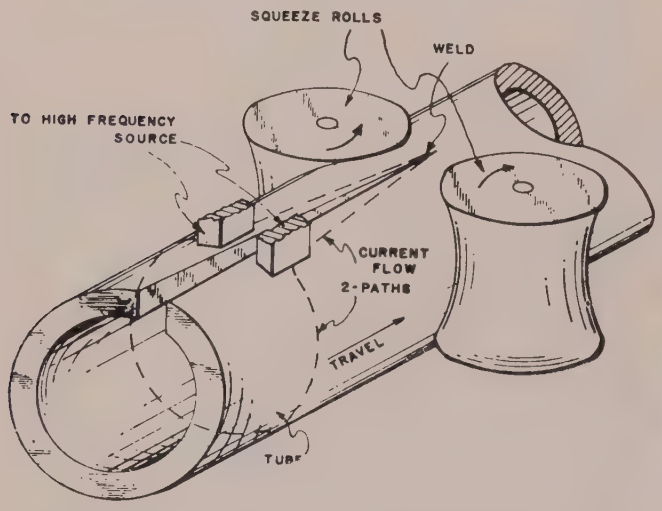


Fig. 5 — Diagrammatic sketch of high-frequency resistance welding equipment for tube welding (courtesy: New Rochelle Tool Corp.)

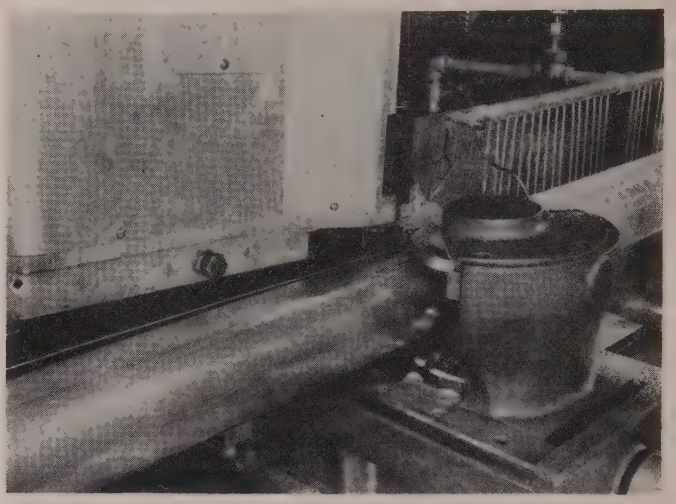


Fig. 6 — Fabrication of 6-in. diameter aluminum tubing with high-frequency resistance welding process (courtesy: New Rochelle Tool Corp.)

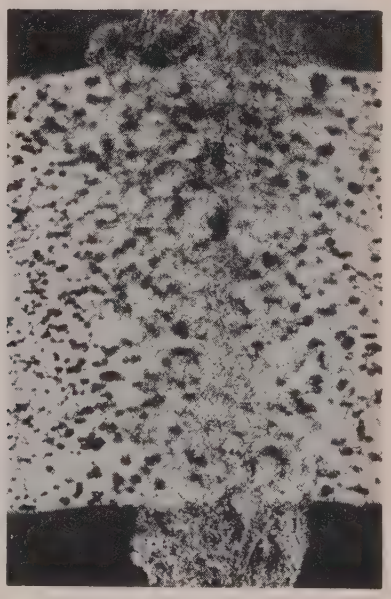


Fig. 7 — Nital etch. Photomicrograph of fusion area of longitudinal weld in 1-in. OD low carbon tubing, welded with the high-frequency resistance welding process. Original magnified 100 times

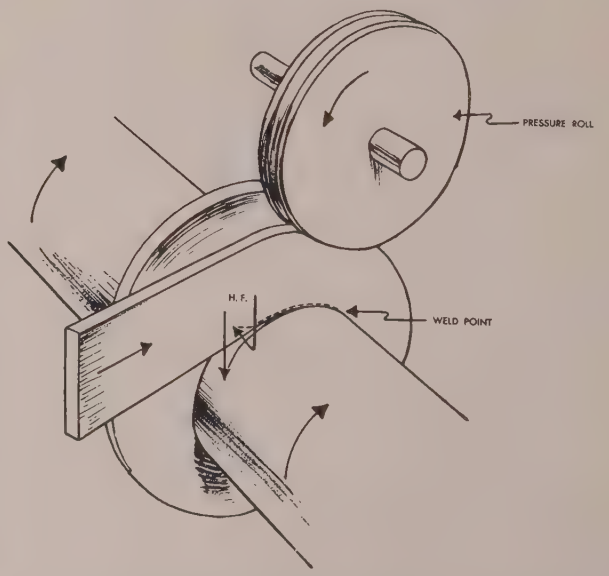


Fig. 8 — Diagrammatic sketch showing method of welding spiral fins to tube with high-frequency resistance welding process (courtesy: New Rochelle Tool Corp.)

quency resistance welding results in greater speeds, wider versatility, and has great potentials for applications where high-production, high-speed welding is important.

Foil Seam Welding

This is a new welding process that was developed in Europe, and has found application where high-speed welding and freedom from distortion is important. This is a resistance welding process for making butt-type joints in flat strips or sheets, using a foil on both top and bottom surfaces over the butt joint. The thin foil strips which are from 5/32–3/16 in. wide and 0.007–0.010 in. in thickness, are used directly over the square abutting edges (Fig. 9).

In this welding operation a small nugget is formed which gradually increases in width (Fig. 9, section CC and DD). The thin foil strip is pressure welded into the fused area, as shown in section DD. A macro of a cross-section of a weld in low carbon 0.090-in. stock is shown in Fig. 10.

An advantage of this process is the high speed of welding, with limited speed from 3–24 fpm on 0.087-in. stock. Another advantage is the freedom from distortion, with the joints being flat and requiring little if any final dressing of the weld metal. The surface of the weld can be painted or glass lined without subsequent grinding or processing operations.

Typical foil seam welding equipment is shown in Fig. 11. Basically the same principles are employed for foil seam welding as for conventional seam welding, execpt for modifications for foil addition. This process has future potentialities because of lower

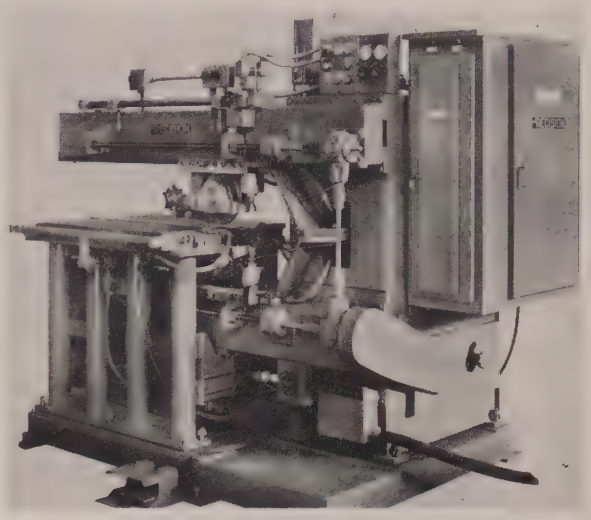


Fig. 11 — Equipment for foil seam resistance welding process (courtesy: Precision Welder and Flexopress Corp.)

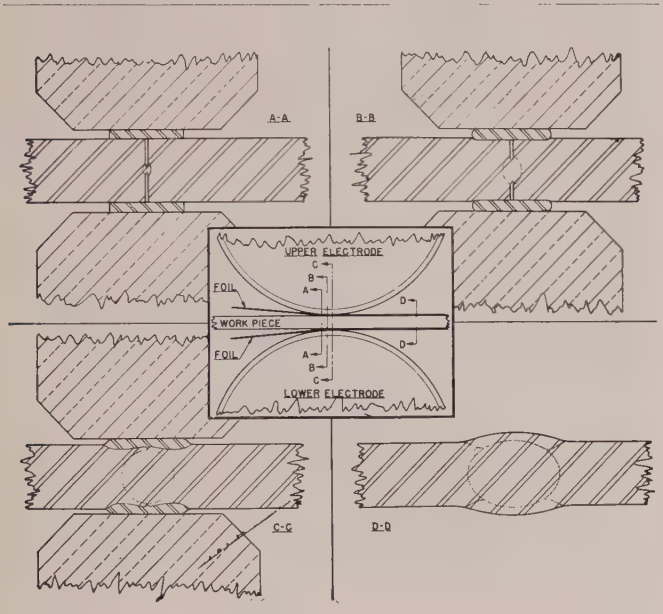


Fig. 9 — Sketches showing position of component parts and progress of fusion with foil seam resistance welding process (courtesy: Precision Welder and Flexopress Corp.)

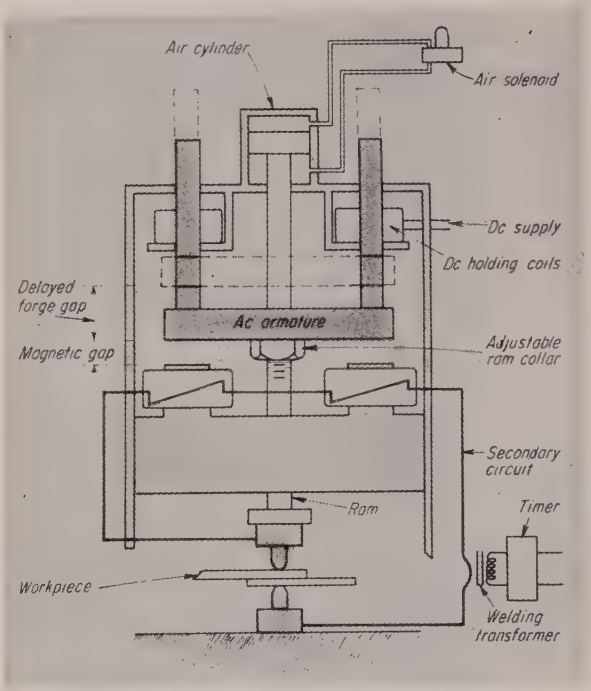


Fig. 12 — Diagrammatic sketch of magnetic force welder (courtesy: Precision Welder and Flexopress Corp.)

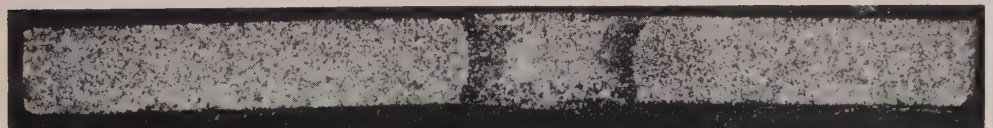


Fig. 10 — Macrograph of typical cross-section of weld made in 0.090 in. thick low carbon steel with the foil seam welding process. Magnified five times

finishing costs and dimensional uniformity.

Magnetic Force Welding

The basic principle of magnetic force welding is the synchronization of the force or pressure by means of coupling this force with magnetic force, created by the welding current. This new principle of force application allows extremely accurate synchronization of the electrode force and the welding current. Resistance welding machines in which the electrode pressure is produced by the electromagnet, which in turn is energized by the welding current, have been under development since 1946. Many of the present difficult welding problems have been satisfactorily solved by this welding technique, where the electrode force is inherently synchronized with the current magnitude and duration.

The magnetic force welder is generally of conventional design and is of a direct-acting piston or ram-type spotwelder, which has been redesigned to accommodate electromagnets for the application of pressure (Fig. 12). The air cylinder actuates the ram for positioning of the weldments between the electrodes, and also imposes a very slight or initial pressure on the work prior to welding. Final force of the weld is controlled by the gap set by the ram collar. Some combinations of materials need a longer heating period before welding, and for this a delay magnet d-c power is mounted as shown. This gives a delayed forge pressure application, which is shown by curves in Fig. 13. As the welding current begins to flow magnetic force is delayed to actuate the forging blow. In most of the applications for this type of weld, the welding time is of short duration, varying from one half to a few cycles.

The advantage of the magnetic force welding process is that it can be used for the welding of high conductivity metals to each other and also to carbon steels. With this process improvement is obtained in the uniformity and consistency of welds being produced, resulting in improved electrical characteristics of contact-type joints.

Many materials previously considered unweldable can now be successfully welded by the magnetic force welder. In the manufacture of electrical contact points, expensive bimetallic contact points can be substituted by less expensive points and joined by magnetic force welding, thereby reducing costs up to 50%. Fig. 14 shows a commercial magnetic force welder.

The principle of magnetic force resistance welding appears to have wide future application to many other resistance welding processes, such as projection welding.

Percussion Welding

Percussion welding is a newly developed resistance welding process in which the weld is made in an exceedingly short period of time over the entire area of the metals to be joined.

The weld can be produced either by a rapid discharge of electrical energy from a stored energy type of power source, or by the rapid dissipation of current with a standard 60-cycle alternating current source. The principle of operation in this process is that the part to be welded has a small pinpoint type of projection formed or coined into the part, which projection is instantaneously vaporized upon contact during the downward movement of the electrode. In the stored energy type of equipment, an air cylinder controls the rapid movement of the electrodes. In the resistance spotwelder, the magnetic force created by the welding current is used in the rapid movement of the electrode, and the welding time is usually one half cycle in duration.

One of the advantages of this process is the welding of large contact areas with the smaller capacity machines and the high quality of weld that is produced. Another advantage is the welding of dissimilar metals for high conductivity electrical contact points, with resultant improvements in finish and appearance. This welding process may have wider applications in the fabrication of ferrous as well as nonferrous metal fields.

Friction Welding

Friction welding is a pressure welding process which was developed in Russia and can be considered as a new addition to the list of new processes. The principle of the operation is based on heat produced by friction of rapidly moving parts under high pressure. The heat obtained is sufficient to cause fusion or bonding of the metal parts.

In the initial tests in the development of this process, the parts were held and rotated in a lathe that was modified to incorporate the application of pressure. Special welding equipment was designed and is shown diagrammatically in Fig. 15. Chucks hold the parts in line with the horizontal axis, one part rotating while the other is held stationary. Pressure is applied which generates sufficient heat so that fusion of the parts takes place. After the temperature is reached, rotation is stopped with the pressure being maintained or sometimes increased until the completion of welding, after which the weld is cooled in air. Total time for some of these parts may be 10–25 sec, and the energy that is consumed may be only 10% of that required for a similar weld fabricated by flash welding.

In this process, foreign materials such as scale and impurities are removed due to friction and plastic deformation of the joined pieces. Metal is plastically deformed and forms a reinforcement as a collar or bead, as shown in Fig. 16. A variety of materials have been reported to have been welded—such as low carbon, medium carbon, free machining

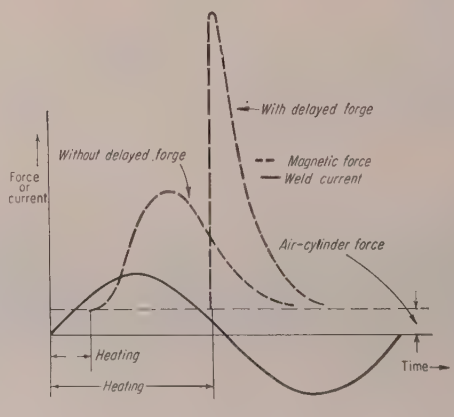


Fig. 13 — Wave form curves showing relation of forging pressure with and without time delay incorporated in magnetic force welder (courtesy: Precision Welder and Flexopress Corp.)

steels, stainless steels, cast iron, brass, aluminum, and titanium, in addition to dissimilar metals such as brass to steel, brass to cast iron, and cast iron to copper.

Limitation of this process at the present time is to parts that are cylindrical in shape and that can be rotated. Fig. 16 shows assemblies of some of the parts which have been fabricated by friction welding in Russia. The two bars shown in the upper part of the view were butted along their axis, with one bar being held fixed while the other rotated at 3000 rpm. The pressure which was applied was approximately 4600 psi. Pressure for other metals such as cast iron would be increased to 9200 psi, and for aluminum it would be reduced to 1120 psi. The heat which is generated is localized, and it may take several seconds to generate the necessary heat.

Some inherent advantages of the friction welding process may be found in low cost of power requirements and high speed of operation. It has interesting possibilities for special applications to production of simple sections, because of lower initial capital investment costs.

Thermo Pressure Welding

Thermo pressure welding is a process that can be considered in the field of solid phase welding. In solid phase welding the parts to be welded are under continuous pressure and the metal is heated until complete or partial plastic flow is obtained.

The success of this process depends upon the cleanliness of the surfaces held under pressure, and the application of heat to produce a true pressure weld. With the thermo pressure welding process it is possible to weld metals to semi-conductors and to other metals which would be difficult to weld by the fusion welding process.

A diagrammatic sketch of parts to be welded by this process is shown in Fig. 17. The sketch shows the attachment welding of filament wire to a heated semi-conductor by the application of pressure. Pressure may range from 5000 to 10,000 psi and the temperature of heating may range from 200–300 C.

In this process the duration of pressure is important, and the time at pressure may vary from 5 sec to 10 min. This welding process has found applications particularly in the joining of semiconductors to metals like germanium and silicon. This process may solve many of the problems related to making

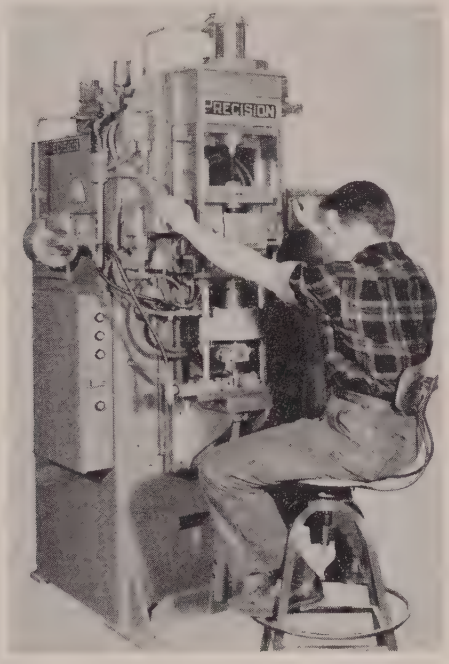


Fig. 14 — Magnetic force welder for commercial application (courtesy: Precision Welder and Flexopress Corp.)

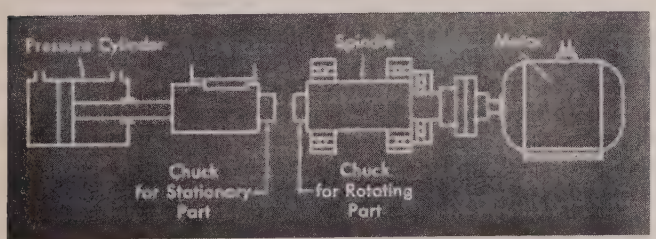


Fig. 15 — Diagrammatic sketch of equipment for friction welding (from Metal Progress, January, 1959, A. B. Tesmen)

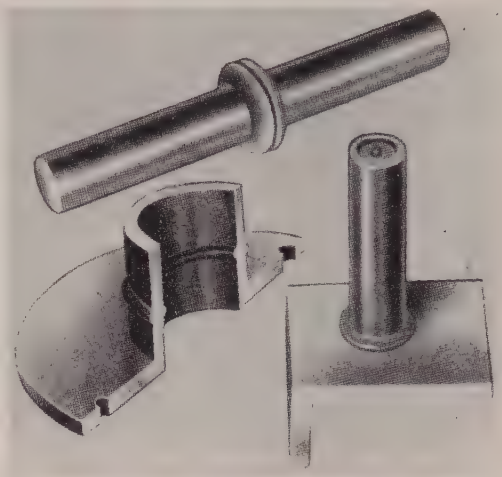


Fig. 16 — Typical parts fabricated by friction welding process (from Metal Progress, January, 1959, A. B. Tesmen)

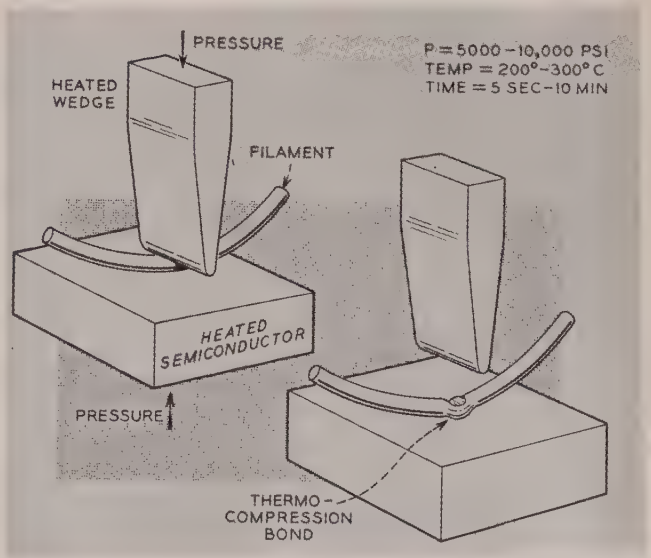


Fig. 17 — Sketches showing position of parts for thermo pressure welding (from Bell Laboratory Record, November, 1957, O. L. Anderson)

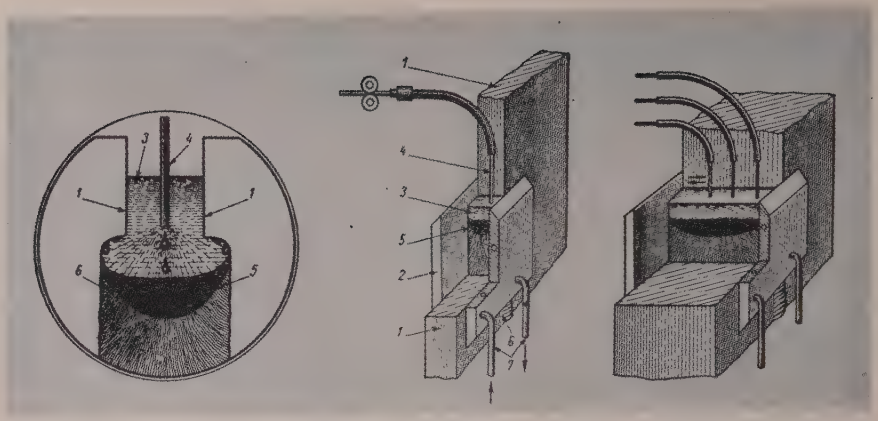


Fig. 18 — Sketches illustrating principles of electro-slag welding process (from Welding Journal, February, 1959)

contacts to small areas without using chemical fluxes and melting in the fabrication of transistors and other semi-conductor devices.

Diffusion Bonding

Diffusion bonding may be defined as a joining method in which metals to be joined are heated to temperatures below their solidus temperature and are held under pressure for some length of time. No liquid phase is formed and the bonding occurs by movement of atoms or intermetallic compounds across the interface on the surfaces, where intimate contact occurs.

At the present time diffusion bonding has been investigated only for certain types of metal combinations, such as: gold and copper, gold and silver, nickel and copper, aluminum and gold, gold and nickel, iron and nickel.

In the diffusion bonding operation intermetallic compounds are formed at the diffusion zone occurring between the metals to be joined. For some of the diffusion bonding operations a plating of nickel, gold, copper, and other metals is first deposited on the surface to be joined. After the plating operation the parts are joined by applying pressure and temperature. Diffusion bonding temperatures are generally below 1000 F and the pressure that is used may range from 4000 to 8000 psi. The duration of time for diffusion bonding varies, depending on the metals to be joined, and may extend from 20 min to 30 hr. During the bonding operation it also may be necessary to use inert gas to prevent oxidation of the bonding surfaces.

In the diffusion bonding operation it is important to maintain cleanliness of the interface surfaces. These surfaces must be carefully prepared to obtain a smooth and uniform intimate contact.

Diffusion bonding may find some important applications in the welding of special alloys where conventional welding processes are not suitable. Since the temperatures are low, previous heat-treatment may not be altered, and also distortion and cleanliness can be controlled. It is anticipated that with further information on the fundamentals of diffusion bonding, a wider application may be created.

Fusion Welding Processes

Fusion welding can be defined as a group of related processes characterized by the use of heat as being of primary importance in obtaining the coalescence of metal. The source of heat may be from an electric power, a chemical reaction, or the combustion of gases and solid fuels. A list of the various fusion welding processes is shown in Table 1.

A number of new processes in fusion welding have also been developed within the last 10 years, which appear to have promise for future applications.

Electro-Slag Welding

One of the most recent developments in fusion welding is electro-slag welding, reported in Russia. Fusion of the electrode wire and the parent metal takes place in and under a molten slag bath covering the metal bath from which the weld is formed. The electrode wire or wires are immersed in the liquid slag. This process is used primarily for welding of heavy sections, and can be used for welding plates ranging in thickness of 2–12 in. or more.

It can be most readily adapted for welding vertical butt joints in heavy weldments, but also can be used for welding of girth seams in heavy wall cylinders, as well as building up large surfaces with weld deposit. Typical electro-slag welding setup is shown in Fig. 18, which illustrates the arrangement of wires in making a vertical weld.

The basic feature of the process is shown in sketch A of Fig. 18. This drawing represents the cross-section of a weld in which a single electrode wire is used. The arc is initially established under the flux, which melts the flux and raises its temperature to a point when the arc action is terminated and the current is conducted through the molten flux. The resistance of the molten flux to the flow of current creates the necessary heat to melt the wire. The weld metal is obtained from the melting of the electrode wire in the superheated flux, and is deposited in a molten puddle of metal, which is directly below the molten flux. The weld is larger and wider than the width of the original component plate edges, as shown in the sketch.

For making vertical butt joints in heavy plates, a number of electrode wires can be used simultaneously, as shown in C of Fig. 18. In some applications the wire can be oscillated back and forth so that the weld is uniformly deposited throughout the entire groove section. For making the vertical butt joints, the ends of the groove are enclosed by water-cooled copper plates so that the slag is confined in the weld area. These copper plates are moved upwards as the deposition of weld metal progresses.

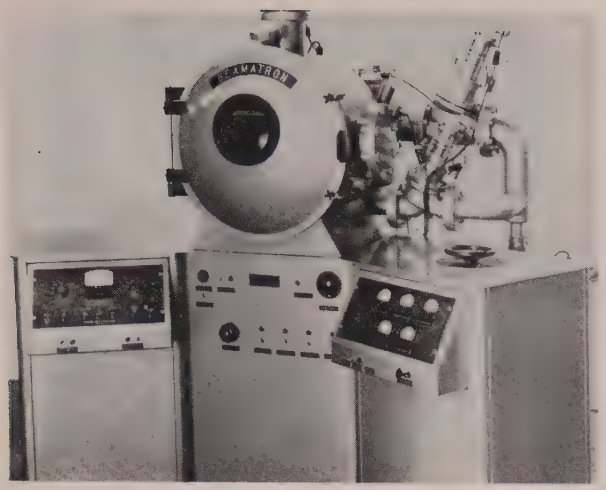


Fig. 19 — Commercial electron beam welding equipment (courtesy: High Vacuum Equipment Corp.)

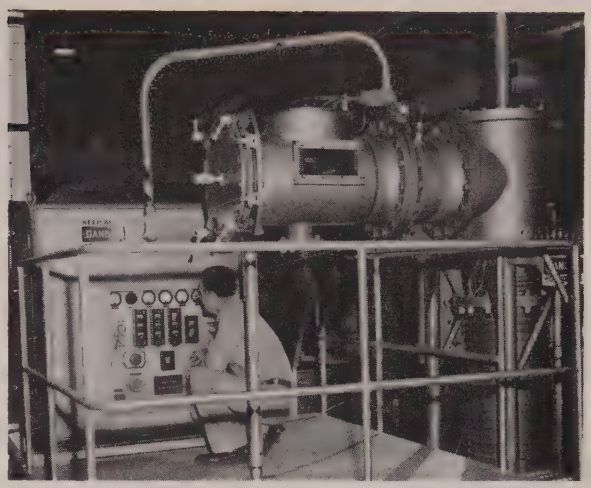


Fig. 20 — Commercial electron beam welding equipment (courtesy: Air Reduction Co.)

One of the advantages of this process is the high metal deposition rate which may be as high as 45 lb of weld metal per hour. Another advantage is the higher thermal efficiency that is obtained with this resistance heating process, when compared to the true arc welding efficiency of the submerged arc welding process.

It is claimed that carbon steels, certain types of low-alloy high-strength steels, and stainless steels have been welded with this process in Russia.

One limitation of this process is the large heat affected zone area adjoining the weld, and the coarse columnar structure of the weld metal deposit. Recommendations have been made to heat-treat the weldments by a normalizing type of postheat-treatments to refine the grain structure. On very large weldments, this would be a costly heat treating operation, and problems of distortion would have to be considered. The process is still so new in this country that the application for industrial use may require further investigation. This process may have definite application for making butt joints in large weldments, such as presses, and also for welding buildups on large surfaces.

Fig. 21 — Internal view of high vacuum chamber used for electron beam welding (courtesy: Air Reduction Co.)

Electron Beam Welding

This is a relatively new welding process, and the principle of its operation is based on the impingement of electrons on the work piece in high vacuum. The electrons are emitted from a tungsten wire filament, and are given a high degree of acceleration and velocity by a high voltage potential between the cathode and anode, the electrons moving from the cathode to the anode. The work piece is the anode.

In order to control the location of the electron beam, it is necessary to concentrate this beam by either magnetic focusing or by electro-static focusing devices. The part to be welded is on a work table or stage, which can be moved so that the welding beam can be directed to the welding area.

Electron beam welding equipment is being manufactured by several companies at the present time, and tests have been made and results reported in various laboratories. Fig. 19 is a unit developed by the High Vacuum Equipment Co. and is called the "Beamatron." The equipment consists of the electron beam gun, a welding chamber, and a vacuum system. It is necessary to reduce the vacuum to a very low value — approximately one-tenth of a micron pressure. The electron beam welder developed by Air Reduction is shown in Fig. 20. Fig. 21 is an internal view of the high-vacuum chamber in which materials are welded with the Air Reduction electron beam welder. Electron beam welders have also been developed by the Carl Zeiss Optical Co. of Western Germany, and the Stauffer-Temescal Corp. of Richmond, Calif. The principles of the high-vacuum electron beam welder as reported by W. L. Wyman of General Electric Co., are schematically shown in Fig. 22. Basically, all of the various electron beam welders operate similarly, the exception being the focusing features of the equipment, which may vary depending on whether the electron beam is directed by means of magnetic or electrostatic controls.

Because a high degree of vacuum is necessary to obtain electron bombardment, the operation is safe from the standpoint of atmospheric contamination. There is no contamination from the oxygen or nitrogen of the atmosphere and from other foreign materials. Since the operation from a purely electronic point of view is possible only below 2×10^{-4} mm

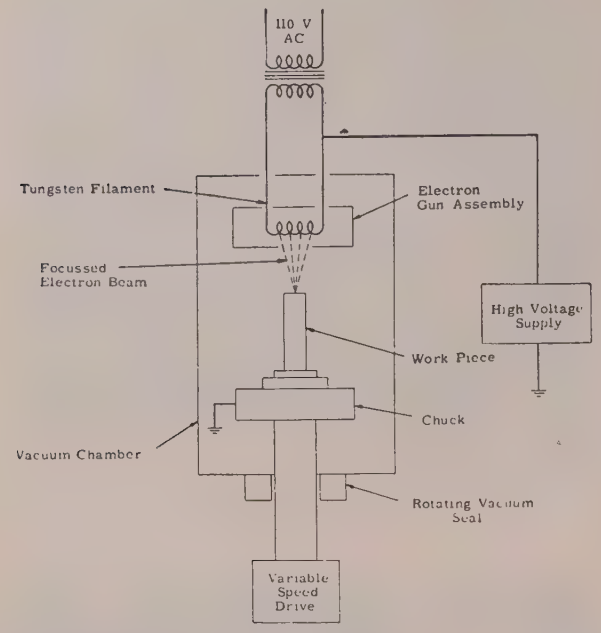


Fig. 22 — Diagrammatic sketch of principles of high vacuum electron beam welder (from Welding Journal, February, 1958, W. L. Wyman, G. E.)

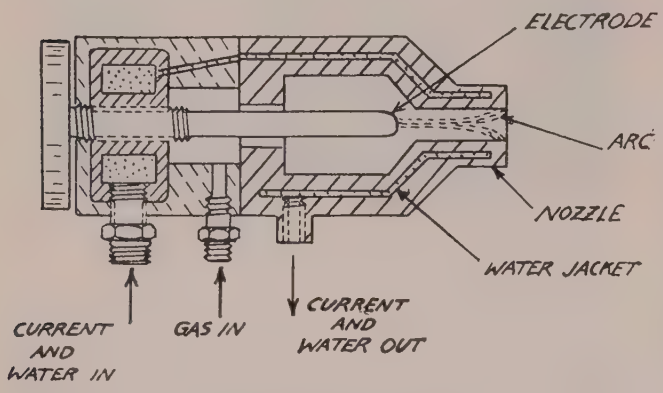


Fig. 23 — Diagrammatic sketch of arc plasma torch (from Welding Engineer, February, 1959)

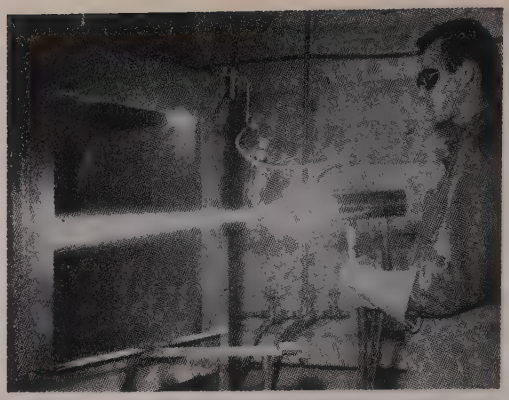


Fig. 24 — Arc plasma torch used for experimental spraying of refractory materials (from Welding Engineer, February, 1959)

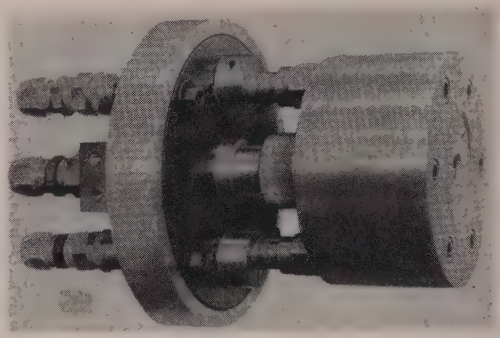


Fig. 25 — Arc plasma gun for commercial applications (courtesy: Plasmatron Co.)

of mercury, the level of atmospheric impurities is less than one part per million. The best commercially available inert gases contain a level of 15–20 parts per million of contaminating impurities. Such high purity is one of the most distinctive advantages of this process over any other welding method that has been developed so far.

In the operation of the electron beam welding, the high temperatures which are developed vaporize some of the impurities in the parent material that is fused, with the resultant weld being much cleaner than the original base material. One advantage is the absence of an electrode or an arc in close proximity to the welding area, which sometimes is a source of contamination since the heating is being done by a beam of electrons originating at some distance from the weld. It is also possible to build into these electron beam welders an optical system for direct viewing, or an indirect optical system that can be focused to the work area. Another significant advantage of this welding process is that the most reactive and refractory metals can be welded without contamination.

Scientific investigation can be made now for the first time to study the effects of welding temperature and the alloying effects of fusion welding without encountering contamination, either by gases or foreign materials. One of the outstanding characteristics of the process is the deep penetration, and tests so far have indicated that it is possible to obtain a depth to fusion width ratio of 2/1, which implies a high-temperature gradient.

Since the process depends upon a high degree of vacuum for its successful application, it is a foolproof method because it will not operate if the vacuum is not maintained at the proper limits. Reports by various investigators have indicated that the quality of the welds is excellent, with freedom from porosity and slag inclusions. Mechanical strength and impact properties are claimed to be comparable or slightly better than welds made by the inert gas tungsten arc process. It is also claimed that the grain size in the fusion zone and the adjacent heat-affected areas is smaller for the electron beam welds than for welds made by other fusion processes.

The following metals have been satisfactorily welded — tungsten, molybdenum, beryllium, tantalum, zirconium, and hafnium. The process can also be applied to some of the more common alloys such as stainless steel and carbon steel. Under certain

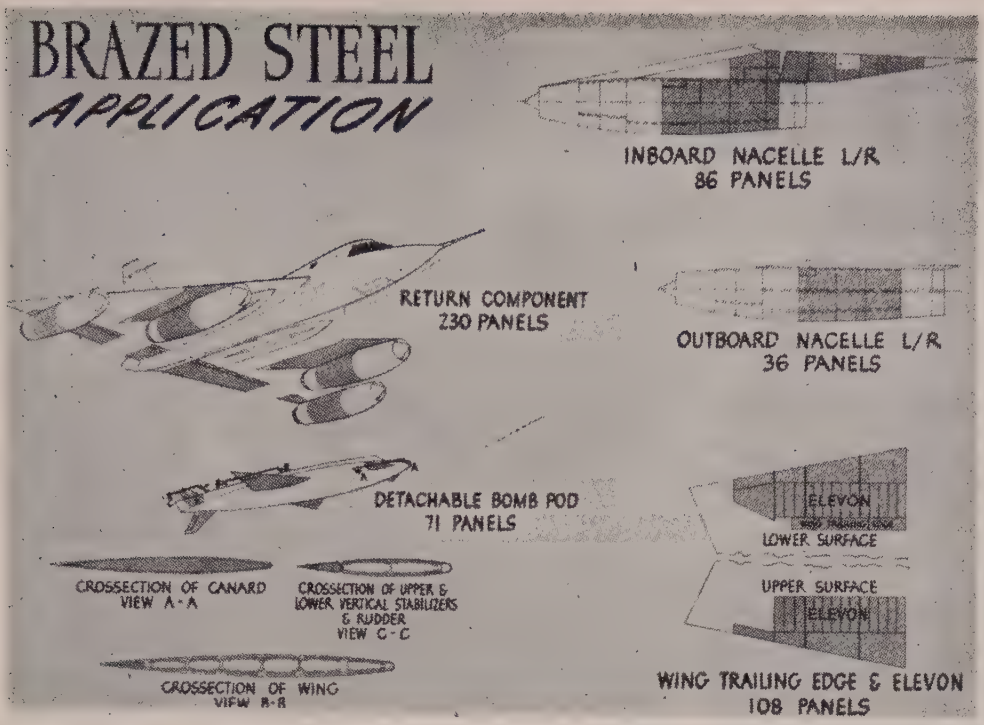


Fig. 26 — Location of brazed sandwich panels, shown in shaded areas, on typical aircraft (courtesy: Convair)

conditions this process may be more economical than other conventional processes. One of its disadvantages when operating at the very high voltages is the emission of penetrating short wave radiation (x-rays), which requires protection or shielding for personnel.

The electron beam circuits are somewhat involved and complicated and require expert technicians to regulate and make the necessary adjustments properly. The equipment must be made vacuum tight for very low pressures, and all openings must be carefully designated to maintain the high vacuum. Since welding is performed in a high vacuum, the welding of large structures would require special large installations and fixtures. However, theoretically there should be no serious limitations for this process. With the increased use of the reactive metals, this process should find wide applications and further developments can be expected.

Arc Plasma

One of the newer achievements in recent years has been the development of arc plasma type of heating, which produces extremely high temperatures ranging from 10,000 to 30,000 F. The basic principle consists of passing a gas — usually the inert gases, although nitrogen and hydrogen can also be used — through a chamber in which an arc is maintained. The flow of electrons from the cathode to the anode ionizes the gas and the resultant bombardment causes the gas to be heated to extremely high temperatures. The passages in the arc torch are constricted so that the walls of the torch, which are water cooled, do not become hot. The ionized gasses are forced through a nozzle opening so that the heated gases have the appearance of a flame emanating from the end of the chamber. Plasma can be defined as any material that is heated to extremely high temperature and is in the ionized condition. It may consist of a hot stream of high-speed gases flowing at speeds from subsonic to ultrasonic ranges. The energy transmitted is largely due to the ionization of the atoms in the vapor and gases, and also to the disassociation of the molecules which results in liberation of large amounts of thermal energy.

Fig. 23 is a diagrammatic view of a plasma flame torch developed by the Thermal Dynamics Corp. The electrode and nozzle are connected through electrical current connectors to a d-c type power source. Plasma forming gas is fed into the nozzle inlet. An electric arc between the electrode and nozzle is struck by super-imposing a high-frequency current, and forms an umbrella shape in the nozzle. Constriction of the arc caused by gas around the electrode point results in formation of an umbrella shape. Gas flows around the electrode and through the umbrella arc, emerging as a plasma flame of 5000–30,000 F. Water flows through the jacket to cool the torch and electric cables are inside water hoses, and hence water-cooled. It is claimed that an overall thermal efficiency of 60–80% can be obtained.

Another application of the plasma torch is for metal spraying of refractory materials. A plasma flame gun is shown in Fig. 24 which is being used experimentally in the spraying of refractory metals. At the present time the arc plasma has been developed primarily as a source of heat, and while some equipment has been developed for metal and ceramic material deposition, the process has a vast potential for future applications for fusion welding of metals. The production of extremely high temperatures is one of the important advantages of the arc plasma process.

Fig. 25 is a photograph of a "Plasmatron" which has been developed by the Plasmatron Corp. Other designs of arc plasma equipment have been developed and investigated which utilize carbon elec-

trodes and water instead of gas as a stabilizing medium. An experimental unit of 10,000-kw capacity has been designed for special high-temperature application.

One limitation of this process is the high-pressure blast effect that is developed in the arc plasma, causing the liquid molten metal to be blown away from the weld area.

Brazing

Brazing can be defined technically as a group of welding processes wherein coalescence is produced by heating to suitable temperatures above 800 F, and by the use of filler metal having a melting point below that of the base metals. The filler metal is distributed between the closely fitted surfaces of the joint by capillary attraction. Flux may or may not be necessary, depending upon the method of heating and metals to be joined.

Brazing is one of the oldest processes for joining metals, utilizing heat and flux, and has found many applications in industry. There are a number of variations of distinct welding processes in the brazing group (as shown in Table 1) and significant improvements have been made, mainly in the brazing alloys and in the heating methods used.

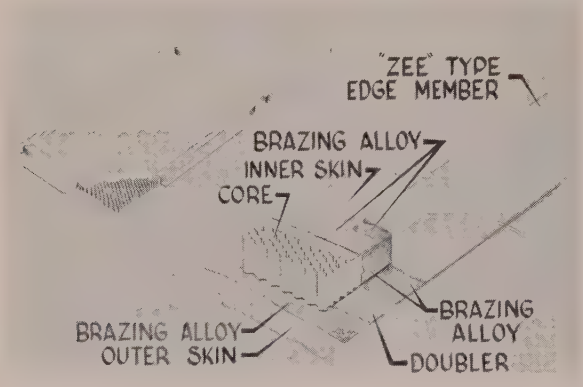


Fig. 27 — Construction of typical flat brazed sandwich panel (courtesy: Convair)

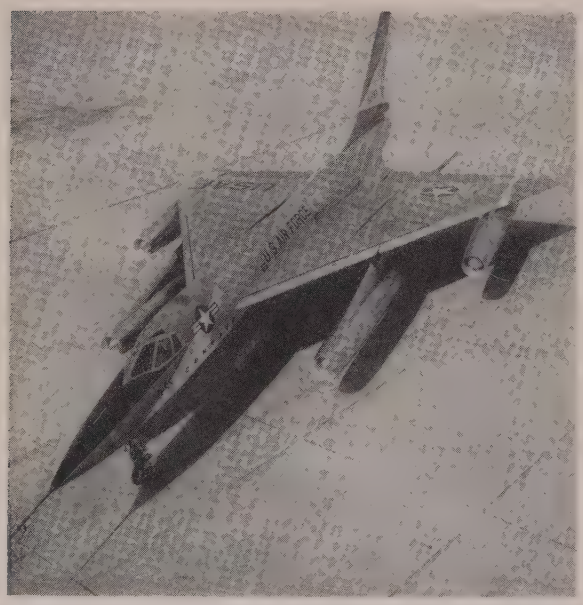


Fig. 28 — B-58 Hustler supersonic bomber utilizing extensive stainless-steel sandwich construction panels (courtesy: Convair)

One outstanding brazing achievement is the development of what is commonly called the "sandwich type construction," which is used on various parts of aircraft components (Fig. 26).

The basic element of brazed sandwich construction is a core of honeycomb or corrugated section made of metal foil which is placed between outer skins, with a suitable brazing alloy interposed between the core and the outer flat surfaces. A thin sheet or foil of brazing alloy is placed between the core and outer surfaces, as shown in Fig. 27. For attachment of such sections to each other, various edge members such as those of the *Z*-type members are also brazed simultaneously in the brazing operation.

The materials can be any of the various types of stainless steels, titanium alloys, and certain high-temperature alloys. The brazing operation can be performed in a number of ways, such as furnace brazing with partial vacuum or an inert gas atmosphere to prevent oxidation, and with graphite blocks to maintain the proper dimensional stability and flatness.

The component parts to be brazed are enclosed in a metal envelope hermetically sealed around the edges, and the entire assembly is placed in a furnace to heat to proper temperature. Other modifications of heating may consist of heating the component parts in a sealed envelope to the proper temperature, and then inserting between two die blocks. Sufficient pressure is exerted to hold and maintain the flatness and shape of the sandwich construction. Excessive pressure must be avoided to prevent buckling. Other forms of heating such as radiant heating, using high-intensity quartz lamps, have lately been investigated and results have been satisfactory.

With respect to brazing alloys, advances have been made in the use of the silver alloys containing small amounts of lithium for improved wetability. Other high-temperature brazing alloys have also been developed, such as the nickel-chromium-boron types.

It is likely that further developments in brazing alloys and procedures will continue, and the applications for joining light gage metals will increase, especially those metals that are difficult to weld by any other process. The future appears promising for marked improvements in design for brazing, as well as improved brazing materials and techniques.

The fabrication of sandwich panels by brazing has been most extensively used in the aircraft industry. In the new B-58 Hustler aircraft, extensive areas have been constructed of sandwich panels. The extent of the use of these sandwich panels is shown in Fig. 28. This construction is used in both the wing and nacelle sections. A typical nacelle section which incorporates sandwich panels is shown in Fig. 29.

The application of the honeycomb sandwich construction is of great importance where the weight-strength ratio is a factor. There is no doubt but that the progress which has been made in the aircraft and guided missile field has been accelerated by the use of the high-strength, lightweight, brazed sandwich construction. The advantage of the braz-

ing process, as applied to honeycomb construction, is that it makes possible the welding or joining of extremely light sections in large areas, for high rigidity and strength. The brazing process is adaptable for the joining of these extremely light sections.

Another advantage is that large areas on panels can be brazed in one operation. In some of the alloys that have been selected, the heat-treatment encountered in brazing can be slightly modified so as to improve or increase the mechanical properties of the alloy. For example, the brazing temperature can be considered as a solution heat-treatment which is followed by a subnormal cooling and a tempering or precipitation treatment.

Disadvantages of the brazing process as applied to sandwich construction are the limitation on the overall size of panels to be welded, the problem of joining panel sections together, and the problem of inspection. Progress has been made in these fields, and in some of the inspection methods, particularly developed for sandwich construction, tests have been established for evaluating the quality of the brazed construction.

Adhesive Bonding

Adhesive bonding can be defined as a process of joining metals by the use of structural adhesives between the metal surfaces. Present adhesives may develop a tensile strength ranging from 7000 to 8000 psi, and shear strengths of 3500–6000 psi.

The structural adhesives used may consist of the epoxy resin type and the synthetic elastomers. In the use of structural adhesives, it is important to prepare the metal surfaces adequately for bonding. In order to obtain the ultimate performance with adhesive bonding, a chemically clean surface is necessary. Where requirements for strength are not severe, vapor degreasing or solvent wiping may be considered. In order to obtain the maximum strength of bonding, the adhesive must physically wet the metal surfaces. Metal preparation can be listed as chemical cleaning, sandblasting, vapor degreasing, and solvent wiping. The first two — chemical cleaning and sandblasting — are preferred for the ultimate performance of the structural adhesives.

In the bonding technique, where the adhesive is of the elastomer type, it is necessary to heat in the range of 200–400 F with sufficient pressure to affect a bond. For the epoxy resin type of adhesives, the temperature is in the range of 300–450 F. Very little or no pressure is required other than that necessary to maintain the parts in contact with each other.

Adhesive bonding is widely used for fabrication of sandwich construction which consists of a core material covered with flat sheets. The core material may be of thin foil aluminum or other metals. The thin foil sheet joining the core structure may be of similar metals. When the edges of the honeycomb cells are bonded to the facing, a lightweight rigid structure results. However, the use of adhesives is not limited to the honeycomb structure, but can also be used on stiffeners, braces, and lap joints.

Adhesive bonding provides continuous contact between mating surfaces, and thus makes possible a seal against the passage of fluids. Adhesive bonding also produces smoother contours and eliminates gaps and bumps which are common to mechanical fasteners. There are no external projections such as rivets or bolt heads, and no surface imperfections such as those obtained in spotwelding operations.

Adhesive bonding has advantages over other methods, since it does not alter the previous heat-treatment given to the metal, and because the construction is continuous. A typical example of adhesive bonding is shown in Fig. 30 which shows a special beaded construction developed by Convair, using structural adhesives for the fabrication of panels.

Fig. 31 is a sketch showing the extent of the adhesive bonded area and the brazed sandwich areas.

Another example of adhesive bonding is the fabrication of helicopter rotor blades. It has been reported that the fly-life of a riveted blade is 90 hr. With a redesign for adhesive bonding, the life expectancy was boosted to 1200 hr. This improvement

Fig. 29 — Aft section of engine nacelles constructed of contoured sandwich panels (courtesy: Convair)

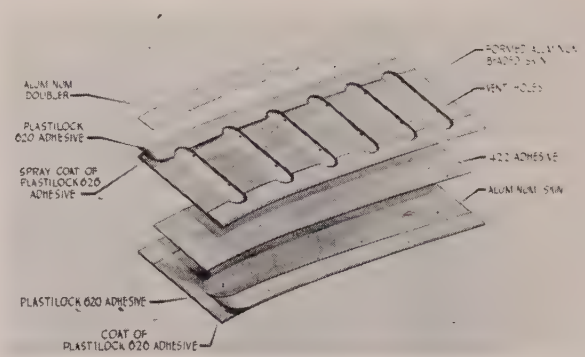


Fig. 30 — Examples of adhesive bonded structure, illustrating special beaded construction (courtesy: Convair)

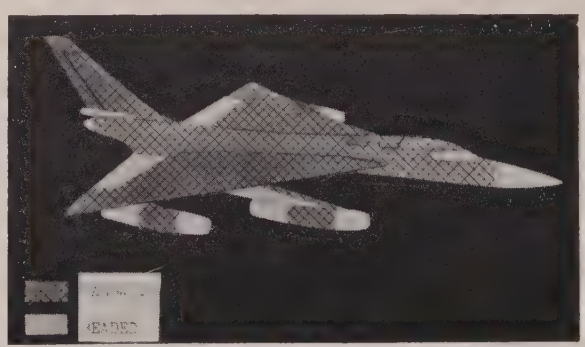


Fig. 31 — Sketch showing extent of adhesive bonded (beaded) area and brazed sandwich construction areas on supersonic aircraft (courtesy: Convair)

was the result of resistance to vibrational fatigue.

Another advantage in using structural adhesives for metal bonding is the distribution of stresses over the entire joint area. In the assembly of heavy gage metal to thin gage by spotwelding or metal fasteners, these processes are limited by the bearing or tearing strength of the thinner sheets. Adhesive bonding does not present localized stress concentration at points of attachment, and full strength of the thin sheets can be realized.

A distinct disadvantage is that the adhesive bonded structures are limited to certain temperature ranges. At the present time the maximum temperature considered is 350 F. Above this requirement the strength of the adhesive bonded section is impaired.

A considerable amount of development work is being performed on structural adhesives which will operate at high temperatures. Tests are under way to raise the operating temepratures from 250–300 F to 450–600 F. This is in experimental stage at the present time.

Another promising adhesive process is the use of glass or ceramics between metal surfaces. Considerable development is being carried out on glass or ceramic materials as a structural adhesive, so that the temperature of operation can be extended above 600 F. Some very promising results have been obtained, and it is expected that further improvements will make it possible to use glass or ceramic structural adhesives, if they may be called such, for aircraft construction.

Summary

In addition to the new welding and adhesive bonding processes that have been described, further developments in the field of automation in joining of metals are to be expected. The future trend will be for more application of automatic welding processes to reduce labor costs.

The field of arc welding by automatic welding controls and welding equipment will be stimulated by the adoption of automatic guiding systems for locating the arc in its proper position. Development in automatic arc welding has reached a point where further application of this process can only be made by a positive method of guidance. Several new automatic guiding methods have been developed. One is the use of a magnetic tape memory type of device that will index the correct position of the welding arc by means of a tracking device, which will direct the arc to the correct position. As the speed of welding is increased, the automatic guiding system and fixturing are necessary requirements.

Another method of automatic guiding is based on the potential differential that exists between the arc and the surrounding metal. A method has been developed which utilizes a minute voltage drop between the arc position and a reference base location in the parent material, so that this small voltage change can be utilized to guide the position of the welding arc. This method has been developed and will undoubtedly have applications for fully automatic arc welding. By automatic guiding, welding can be performed in irregular sections at high speeds.

We can also depend on new important developments in many of the established welding processes such as the metal-arc coated electrode process, where we expect some radical improvements in quality and performance of the covered electrodes. These improvements will be in the low hydrogen type of electrodes, that will be nonhydroscopic and can be used without the present precautions with regard to storage.

Welding power sources will also be radically improved in the future, as to their use in the various arc welding processes. It is also expected that in the field of resistance welding, improvements will be made in the speed and quality of welds and the applications of these processes to welding of dissimilar metals.

The author wishes to acknowledge the valuable assistance given by Ivan Kutuchief, Senior Welding Engineer, A. O. Smith Corp., in the preparation of this presentation.

References

1. "Electro-Slag Welding." *Welding Journal,* Vol. 38, February, 1959, pp. 132–133.

2. "Welding Heavy Sections by Russian Electro-Slag Process," by W. Anders. *Welding Journal,* Vol. 36, June, 1957, pp. 290s, 306s, 311s.

3. "Electric-Slag Welding Process," by Gerald E. Claussen. *Welding Journal,* Vol. 36, August, 1957, pp. 381s–392s.

4. "High Frequency Resistance Welding," by Wallace C. Rudd. *Welding Journal,* Vol. 36, July, 1957, pp. 703–707.

5. "Electron Beam Welding." *American Machinist,* Vol. 103, Feb. 23, 1959, pp. 95–98.

6. "Adhesion of Solids — Principles and Applications," by O. L. Anderson. *Bell Laboratory Record,* Vol. 35, November, 1957.

7. "Epoxy-Resin Adhesives," "Metal to Metal Bonding," "Honeycomb Sandwich Type Materials." Engineering Bulletins, Minnesota Mining and Manufacturing Co., Detroit, Mich., October, 1957.

8. "Metal to Metal Adhesives for Assembly of Aircraft," by Richard G. Newhall. Western Business Publications, San Francisco, 1955.

9. "Brazing Hustler's Stainless Sandwich," by A. A. Lanzara and R. E. Purnell. *Aircraft and Missiles Manufacturing,* Vol. 1, March, 1958, pp. 10–16.

10. "High Vacuum Electron Beam Fusion Welding," by W. L. Wyman. *Welding Journal,* Vol. 37, February, 1958, pp. 49s–53s.

11. "Adhesive Bonding of Metals," by George Epstein. Pub. by Reinhold Book Publishing Co., New York, 1954.

12. "Recent Developments in Magnetic Force Welding," by E. J. Funk, Precision Welder and Flexopress Corp., Cincinnati, Ohio.

13. "Diffusion Bonding Below 1000 F," by J. T. Niemann, R. P. Sopher, and P. J. Rippel. *Welding Journal,* Vol. 37, August, 1958, pp. 337s–342s.

14. "Plasma Jet," by Gabriel M. Giannini. *Scientific American,* Vol. 197, August, 1957, pp. 80–84.

15. "Plasma Heat — Newest Addition to Welding's Product Lineup." *Welding Engineer,* Vol. 44, February, 1959, pp. 50–52.

16. "Ultrasonic Welding—New Technique Grows," by J. Byron Jones. *Metal Progress,* Vol. 73, April, 1958, pp. 68–77.

17. Welding Handbook, 1948. American Welding Society.

HIGH-ENERGY ACCELERATORS

Keith R. Symon
University of Wisconsin and Midwestern Universities Research Assoc.

HIGH-ENERGY particle accelerators are among the most important, most complex, and most expensive research tools now used in scientific research. This paper presents a summary of the subject to date.

To accelerate atomic particles to energies of billions of volts, it is necessary to provide some means whereby the particles can be given successive kicks in energy — the total of the energy received in all of these kicks adding up to the desired final energy. Such high-energy accelerators can be classified into two categories: linear and circular. The author discusses the advantages and disadvantages of both.*

HIGH-ENERGY accelerators are among the most important, most complex, and most expensive research tools now used in scientific research. In expense and complexity they are perhaps rivaled only by earth satellites. Therefore, it seems advisable to begin this discussion of accelerators with the question, what are high-energy accelerators used for and why do we need them?

High-energy physics is a new area of physics which lies at the frontier of present-day scientific research. During the years from 1900 until about 1930, atomic physics was at the scientific frontier. During that time, we were learning about the electronic structure of atoms and molecules. From 1920 until about 1945, the frontier had moved into the field of nuclear physics. Attention was now focused on the structure of the nucleus of the atom itself. We have now come a rather long way toward unraveling the structure of the atomic nucleus and learning the nature of the forces which hold together the protons and neutrons making up the nucleus. Since about 1945, attention has begun to shift toward the problem of the nature and the structure of the fundamental physical particles themselves.[1] We would now like to know how a proton or a neutron or an electron or any other particle is constructed and what are the fundamental laws governing its behavior. Thus, the history of physical research in the present century reveals a shifting focus of interest into the nature of smaller and smaller physical structures.

The exploration of ever smaller physical structures requires the use of probing particles of ever higher energies. To see why this is true, we may refer to Eq. 1 which represents the fundamental formula of quantum mechanics first enunciated by

* Paper presented at SAE Summer Meeting, Atlantic City, June 15, 1959.
[1] "Elementary Particles," by M. Gell-Mann. *Scientific American*, Vol. 197, July, 1957, p. 72.

Max Planck:

$$E = hf = hc/l \quad (1)$$

where:

E = Energy of a light quantum or any other physical particle
f = Corresponding frequency of the light wave
c = Velocity of light
l = Wavelength of the light
h = Planck's constant

If we want to study some physical structure by utilizing light waves, then a fundamental principle of optics tells us that the wavelength of the light used must be shorter than the dimensions of the structure we are trying to study. Since the quantum theory tells us that all particles have a wave nature, the same remarks would apply to an attempt to study a structure by bombarding it with physical particles, as for example, in an electron microscope. In all such cases there is a wavelength l which must be smaller than the dimensions of the structure being studied. From Eq. 1 this implies that the energy of the corresponding quantum must be larger, the smaller the structure being studied. The size of a proton, for example, which is one of the fundamental particles we would like to study, is of the order of one ten trillionth of a centimeter. If we substitute this value for l in Eq. 1, we find that the energy of a light quantum or other particle of wavelength small enough to probe the structure of a proton would have to be more than 1 bev.[2] We use as our unit of energy the energy that an electron receives when it falls through a potential difference of one volt. Another way of arriving at the same result is to consider the famous relation represented by Eq. 2:

$$E = mc^2 \quad (2)$$

where:

m = Mass of a particle
E = Energy equivalent to that amount of mass
c = Speed of light

One way of studying the structure of a proton, for example, would be to try to make one. If we substitute in Eq. 2 the mass of a proton, we find that the amount of energy which would be required to make a proton is again about 1 bev. As a matter of fact, protons can only be made in pairs, a proton and an antiproton, and, therefore, the energy required is 2 bev. Furthermore, when one considers that in an atomic collision a good deal of the incident energy is wasted, it turns out that to make a proton-antiproton pair by letting a high energy proton collide with a stationary one requires an energy of 6 bev. This corresponds to the energy of the Bevatron at the University of California which was the first man-made device capable of producing artificial proton-antiproton pairs.

The earliest research in high-energy physics involved the study of the very rare high-energy events produced by cosmic rays. The construction of high-energy accelerators during the last 10 years has provided us with artificial sources of multibillion volt particles. The study of high-energy physics has revealed a completely new and unexpected realm of physical phenomena associated with the so-called strange particles. In addition to the ordinary particles like the proton, neutron, electron, photon, and the mesons which are associated with nuclear forces, we have discovered new classes of particles, the heavy mesons and hyperons. They are given the technical name "strange particles," because their existence and properties were completely unexpected before their discovery.

Since there are now more than 40 different kinds of so-called fundamental particles known in physics which enter into various kinds of reactions with one another, we are faced with the challenge of learning how to classify these various particles and to account for the various processes which can take place. This area of science is about at the point where chemistry was at the time of the discovery of the periodic table. We now know how to arrange these particles in a table, something like the periodic table of the elements, in such a way that one can predict with some degree of success what kinds of particles exist, and what kinds of reactions will take place between the particles (just as the valence rules in chemistry predict what sorts of chemical reactions may take place). We have, however, as yet almost no theoretical understanding of why the table is constructed as it is. We are not even sure whether the table is yet complete.

To a scientist the challenge to explore the mysteries of particle physics is sufficient justification in itself for doing high-energy research, just as the challenge to explore space would be sufficient justification for doing so regardless of what practical results might come from such exploration. History tells us that important practical applications have in the past always resulted from new knowledge of natural laws. It is impossible at the present time even to imagine what practical application might arise from the knowledge of the properties of the new strange particles which exist for only tiny fractions of a second in high-energy processes. We should, however, also remark that in, say 1930, it would have been impossible to imagine any practical application that might result from a study of the properties of the incredibly tiny atomic nucleus.

Let us turn now to particle accelerators which are the tools needed for research in this fascinating new field. In order to accelerate atomic particles to energies of billions of volts, it is necessary to provide some means whereby the particles can be given successive kicks in energy. The sum total of the energy received in all of these kicks eventually adds up to the desired final energy. We may first classify high-energy accelerators into two general types. In a linear accelerator the particle travels in a straight line through a succession of accelerating gaps. A high radio frequency voltage is applied to each of these gaps which are so arranged that the particle arrives at each gap when the voltage is in such a direction as to accelerate the particle through the gap. In a linear accelerator the particle necessarily passes through each accelerating gap only once. In a circular accelerator the particles are made to travel in circular orbits by means of a magnetic field so that they may pass through the same accelerating gap a great number of times. Linear and circular accelerators each have their own characteristic advantages and disadvantages and both types are important in high-energy research.[3]

[2] Billion electron volts (bev).
[3] "High Energy Accelerators," by M. S. Livingston. Pub. by Interscience Publishers, Inc., New York, 1954.

The highest energy linear accelerator now in existence, which accelerates electrons to energies up to 700 mev is at Stanford University. Nine additional accelerating sections are being added to the present 21 sections in order to extend the energy of the Stanford linear accelerator to over 1 bev. Experiments on the scattering of high-energy electrons by neutrons and protons have already yielded valuable information about the structure of neutrons and protons.

Stanford University has recently proposed to construct a 2-mile-long linear accelerator which would accelerate electrons to energies of from 15 to 45 bev. This accelerator, when constructed, will provide electrons in an energy range well beyond that of any other electron accelerator now in existence or contemplated. It will make it possible to extend the Stanford work on the structure of neutrons and protons well into the multibillion volt range. It will also provide a great deal of new information about strange particles and may even make it possible to produce new kinds of particles not yet known. The design of this 2-mile accelerator represents a fairly direct extension of the design of the 700-mev accelerator now at Stanford. The 2-mile accelerator, however, will need to have about 40 times as many accelerating units.

All existing accelerators for accelerating protons to the billion volt energy range are of the circular type. A great many varieties of circular accelerators are possible. We may first classify circular accelerators according to whether the magnetic field is pulsed or fixed. In a pulsed field accelerator the accelerated protons revolve in an orbit of constant radius and the magnetic field is increased with time as the particles are accelerated so as to maintain a constant orbit radius. In a fixed field accelerator, the magnetic field is constant in time and the accelerated particles move to orbits of larger radius. An example of a fixed field accelerator in the low-energy range is the familiar cyclotron. An example of a pulsed field accelerator in the low-energy range is the betatron. Next, we may classify accelerators according to the kind of focusing which is used in order to make the particles circulate in stable orbits. The earliest circular accelerators constructed have a magnetic field which decreases slightly as we move out along a radius from the center of the accelerator and which is uniform as we move in a circle around the center of the accelerator. This kind of magnetic field pattern is called constant gradient or weak focusing. A new kind of focusing principle was recently discovered by Christophilos in Greece and by Courant, Livingston, and Snyder at Brookhaven.[4] This principle, called alternating gradient or strong focusing, utilizes a series of magnets arranged around the circumference of the circular orbit and so designed that the magnetic field gradient alternates from one magnet to the next. It can be shown that in this way considerably stronger focusing of the orbits may be obtained. This makes it possible to design an accelerator with a smaller aperture within which the particles circulate and with a consequent saving in the weight of the magnets. Finally, we may classify accelerators according to the way in which the particles are accelerated. All high-energy accelerators are synchrotrons, that is, they are accelerators in which a radio frequency voltage is maintained across one or more accelerating gaps around the circumference of the accelerator and in which the frequency of this voltage is maintained in synchronism with the frequency of revolution of the particles as they are accelerated.

The first accelerator in the multibillion volt range was the Cosmotron at the Brookhaven National Laboratory which began operating in the spring of 1952. The Cosmotron is a weak-focusing proton synchrotron which can reach energies up to 3 bev. The highest energy accelerator in the United States at present is the Bevatron at the University of California in Berkeley (Fig. 1). The Bevatron is a 6.2-bev weak-focusing proton synchrotron which has been in operation since 1954. Both the Cosmotron and the Bevatron have been utilized in the study of the properties of strange particles. In addition, a few years ago the first proof of the existence of the antiproton was achieved in experiments at the Bevatron. The highest energy accelerator in the world today is at the Dubna Laboratory outside of Moscow. This is a 10-bev weak-focusing proton synchrotron very similar to the Bevatron. It is housed in the building shown in Fig. 2.

The discovery of the alternating gradient focusing principle allows an increase in the maximum energy accelerator which it is economical to construct. Pulsed field alternating gradient proton synchro-

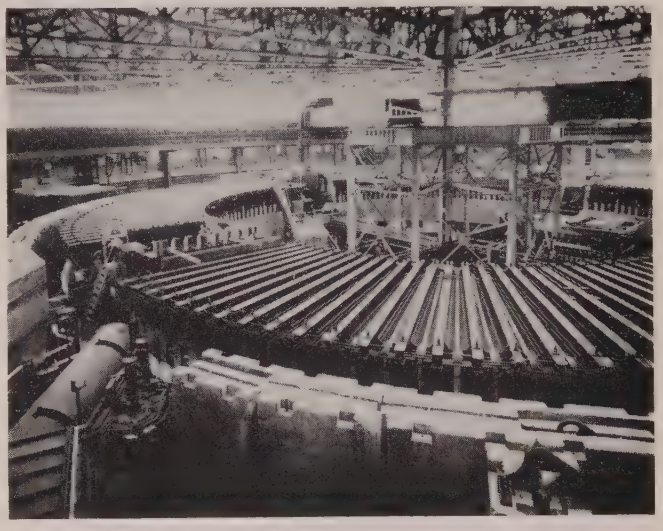


Fig. 1 — The Bevatron (courtesy of Lawrence Radiation Laboratory)

Fig. 2 — Synchrophasotron building (from a Soviet newspaper)

[4] "Strong-Focusing Synchrotron," by E. D. Courant, M. S. Livingston, and H. Snyder. *Physical Review*, Vol. 88, Dec. 1, 1952, pp. 1190.

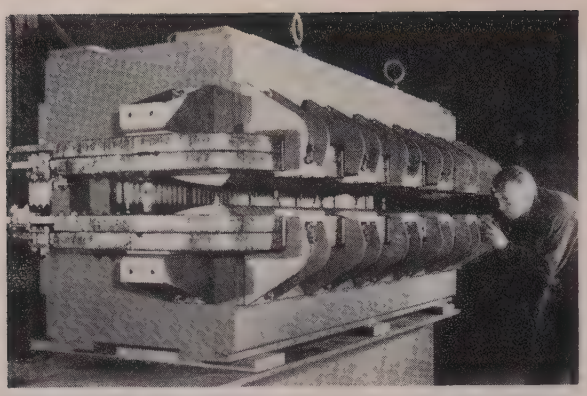


Fig. 3 — Magnet for Brookhaven alternating gradient synchrotron (courtesy of Brookhaven National Laboratory)

Fig. 4 — Brookhaven alternating gradient synchrotron (courtesy of Brookhaven National Laboratory)

trons which will reach energies of 25–30 bev are now under construction at the Brookhaven National Laboratory and at the CERN Laboratory. CERN is a cooperative laboratory set up by 12 European nations near Geneva, Switzerland. Fig. 3 is a picture of one of the 240 magnets for the Brookhaven alternating gradient synchrotron. These magnets will be placed in a circular ring inside a buried tunnel. An aerial view is shown in Fig. 4. Straddling the ring at one side is an experimental building in which experiments with the high-energy proton beams will be performed. Near the experimental building is the laboratory building which houses equipment and the people who work on the accelerator. At the other side of the ring may be seen the injector building which houses a 50-mev linear accelerator. Fifty mev protons are injected into the ring and then accelerated to energies of 25–30 bev. In the upper right-hand corner of the picture can be seen the building which houses the Cosmotron. Fig. 5 is a view inside the experimental building. One can see the magnet foundations along a portion of the circular arc. At the time this picture was taken, the magnets had not yet been installed. One can also see under construction the thick concrete shielding walls which are necessary to protect people and equipment in the building from the high-energy radiation. The Brookhaven alternating gradient synchrotron is scheduled for completion in

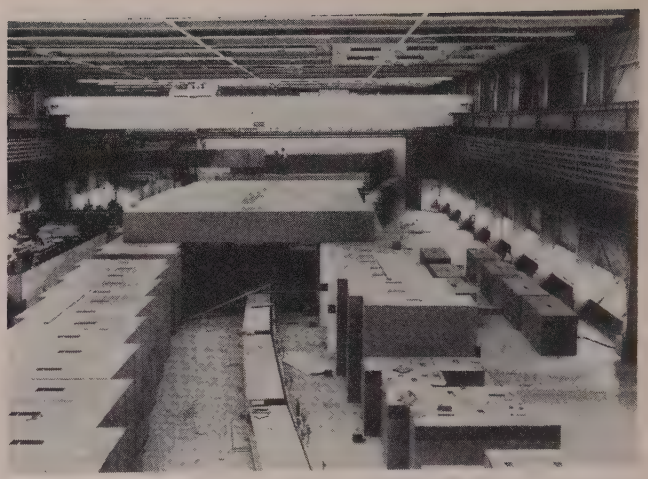


Fig. 5 — Brookhaven alternating gradient synchrotron experimental area (courtesy of Brookhaven National Laboratory)

early 1960.

There are a number of other high-energy accelerators in existence or under construction in the United States including two 1-bev electron synchrotrons, one at Cornell University and the other at the California Institute of Technology, a 3-bev proton synchrotron under construction by the University of Pennsylvania and Princeton University, a 6-bev strong-focusing electron synchrotron under construction by Harvard University and the Massachusetts Institute of Technology and a 12.5-bev weak-focusing proton synchrotron whose design is nearly completed at the Argonne National Laboratory.

The Midwestern Universities Research Association (MURA) laboratory set up by 15 universities in the Midwest has been engaged during the last few years in the development of a new type of accelerator which promises to extend both the beam intensities and the collision energies which can be achieved with high-energy accelerators. This new type is a fixed-field alternating gradient (FFAG) synchrotron.[5] In an FFAG synchrotron particles are accelerated in a donut-shaped vacuum tank. The particles are maintained in their orbits by means of a fixed magnetic field provided by a series of magnets placed around a ring, as indicated in Fig. 6. The magnets are so designed that the magnetic field is weak at the inner edge of the vacuum tank and strong at the outer edge. Thus, low-energy particles can circulate near the inside edge of the vacuum tank while high-energy particles may circulate near the outer edge. In order to provide stable orbits over a wide range of energies, focusing is achieved by utilizing the alternating gradient principle. There are two types of FFAG magnetic field patterns. The radial sector pattern is shown in Fig. 6. In this case, an alternating magnetic field gradient is attained by providing two kinds of magnets, positive and negative magnets, placed in alternation around the ring. In the positive magnets the magnetic field is in such a direction as to curve the particles around the center of the accelerator. In the negative magnets the magnetic field is in the opposite direction. The only purpose

[5] "Fixed-Field Alternating-Gradient Particle Acclerators," by K. R. Symon, D. W. Kerst, L. W. Jones, L. J. Laslett, and K. M. Terwilliger. *Physical Review*, Vol. 103, Sept. 15, 1956, p. 1837.

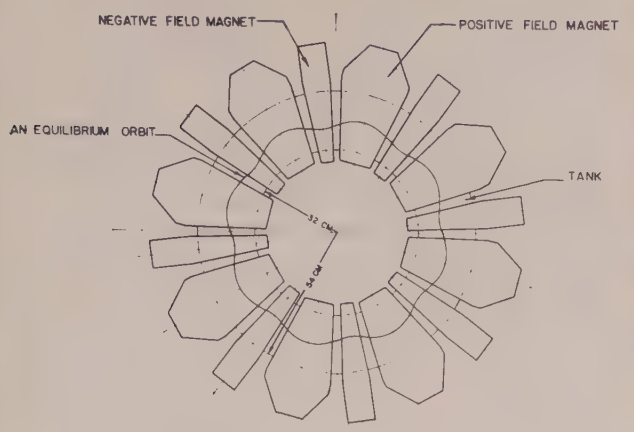


Fig. 6 — Magnet layout with equilibrium orbit, FFAG model

Fig. 7 — Radial sector model electron accelerator, 400 kev (courtesy of Midwestern Universities Research Assoc.)

of the negative magnets is to provide alternating gradient focusing in order to stabilize the orbits. If the negative magnets were not present, the particle orbits would be unstable and the particles would fly out vertically and strike the top and bottom of the vacuum tank. It is easy to see from the figure that a consequence of the insertion of negative magnets is to make the accelerator larger than it would otherwise have to be. Calculations indicate that such an accelerator would have to be three to five times greater in diameter than a more conventional type accelerator at the same energy. Fig. 7 is a picture of a small model radial sector accelerator which was constructed about three years ago in order to study experimentally the properties of an accelerator of this type. The model shown in Fig. 7 will accelerate electrons to an energy of 400,000 volts.

In order to reduce the size of FFAG accelerators, the spiral sector pattern was suggested by D. W. Kerst. A model spiral sector electron accelerator which has been constructed is shown in Fig. 8. The dark areas in the picture are the spiral magnets. If we imagine taking a radial sector accelerator like the one shown in the previous picture, removing the negative magnets and twisting the positive magnets into a spiral, we obtain the spiral sector pattern. This design achieves alternating gradient focusing without the need for negative magnets. Such an accelerator, therefore, need have a diameter only slightly larger than more conventional accelerator types.

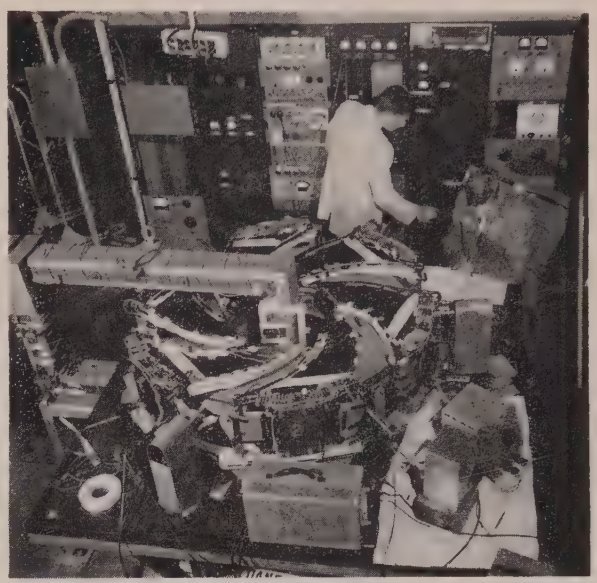


Fig. 8 — Small spiral sector of FFAG accelerator (courtesy of Midwestern University Research Assoc.)

Among the advantages in utilizing fixed-field accelerators, perhaps the most important is the flexibility which is thereby gained in the design of a system to accelerate the particles. Once a suitable magnetic field pattern has been achieved, one may think of the magnetic field as simply a box which is capable of holding particles on stable orbits at all energies from the low energy at which they are injected into the accelerator to the highest energy to which they are to be accelerated. We may then proceed to accelerate them by means of applied radio frequency voltages according to any acceleration schedule which seems most desirable. In particular, it is possible to design acceleration schemes in which many groups of particles can be acceler-

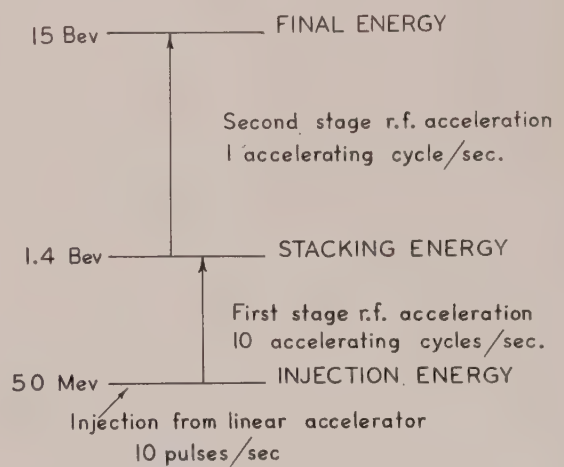


Fig. 9 — Proposed accelerating scheme utilizing beam stacking

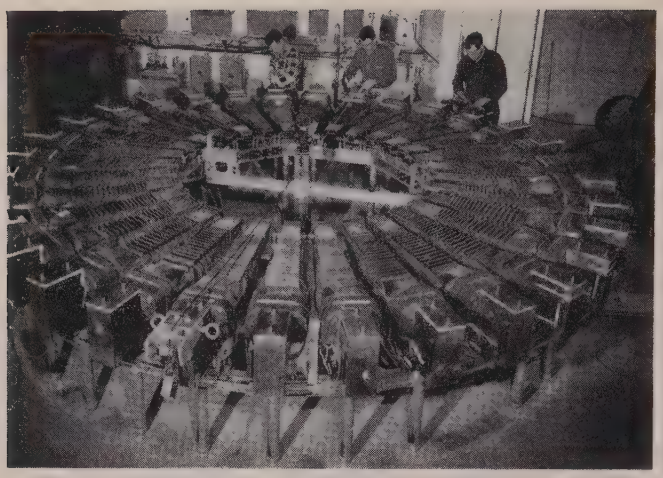


Fig. 10 — Forty-mev two-way electron accelerator (courtesy of Midwestern Universities Research Assoc.)

Table 1 — Energy Relations in Proton-Proton Collisions

Available Energy, bev	Required Incident Beam Energy bev	Required Colliding Beam Energy, bev
1.12	2.90	0.56
2.02	6.20	1.01
5.23	25.0	2.61
7.99	50.0	4.00
30.0	540.0	15.0

ated simultaneously. To take an example, consider the acceleration scheme shown diagrammatically in Fig. 9. This represents a possible acceleration scheme for a 15-bev accelerator. Ten times each second a pulse of protons is injected at 50 mev in to the accelerator and accelerated by means of radio frequency voltages to an energy of 1.4 bev. The accelerated particles are left coasting at 1.4 bev until 10 successive groups of particles have been accelerated to that energy. This process is called "beam stacking." Once each second, a second radio frequency accelerating system picks up 10 groups of particles which have been stacked at 1.4 bev and accelerates them to the final energy of 15 bev. The advantage of this scheme is that many groups of particles can be injected into the accelerator per second. At the same time the second accelerating stage during which the particles receive most of their energy need only operate at a repetition rate of one per second. Repetition rates in existing high energy accelerators are limited by the rate at which the magnets can be pulsed, which is about once every four or five seconds. Thus, the acceleration scheme shown in Fig. 9 for an FFAG accelerator can accelerate 50 or more times as many pulses of particles per second as present-day accelerators. In this way an increase in the intensity of the accelerated beams may be achieved. It should be remarked that it is also possible to design pulsed field accelerators with high repetition rates, as in the case of the Princeton-Pennsylvania and Harvard-M.I.T. accelerators mentioned earlier.

A second possibility opened up for the first time by the use of beam stacking in FFAG accelerators is the possibility of accumulating an intense beam of circulating high-energy particles within the accelerator. If, instead of allowing the particles to strike a target or extracting them from the accelerator, we simply leave the accelerated particles circulating at say 15 bev, then calculations indicate that currents as high as 50 amp of circulating protons can be achieved. It then becomes practical to imagine allowing two such intense circulating beams to intersect in opposite directions so that the oppositely circulating protons may collide.[6] The probability that two protons traveling in opposite directions will collide with one another is extremely small so that heretofore it has been out of the question to imagine letting beams of protons from existing accelerators collide with one another. However, by utilizing stacked beams it is possible to have as many as 10^{14} or 10^{15} protons circulating within an FFAG accelerator. If beams of this intensity are made to collide, we may expect that as many as a hundred thousand or more proton-proton collisions will occur per second, which is an adequate number for experimental purposes.

The advantage of utilizing colliding beams is indicated by Table 1. In order to understand the table, it is necessary to realize that in an experiment of the type conducted with present-day accelerators, a high-energy proton is allowed to collide with a stationary proton in some solid or liquid target. In such a collision, a large fraction of the energy of the incident proton is dissipated in forward motion of the two particles after the collision. Only a fraction of the energy is available for producing new particles. In the first column in the table, we show the available energy in a collision of two protons. It is this energy which determines the nature of the collision which takes place. In the second column, we show the energy required in a proton beam which is allowed to strike a stationary target in order to achieve the available energy shown in the first column. Thus, for example, the Cosmotron produces protons at an energy of 2.9 bev. The available energy in the reaction however is only 1.1 bev. The Bevatron produces protons at an energy of 6.2 bev. The available energy in the reaction is only a little over 2 bev, barely enough to make a proton-antiproton pair. The new alternating gradient synchrotron under construction at Brookhaven will produce an energy of 25 bev but the available energy in a proton-proton collision will only be 5.2 bev. It is readily seen from the table that the energy required of an incident proton increases very rapidly while the available energy in the collision increases only rather slowly. This is a consequence of the theory of relativity, according to which a fast moving particle has a greater mass than when it is at rest. Thus, a 25-bev proton has a mass about 25 times as great as a proton at rest. If we imagine, for example, a fast baseball striking a ping-pong ball a rest, we can easily see that most of the energy of the baseball simply goes into forward motion of the baseball and the ping-pong ball, and very little of it is available for producing noise or denting the ping-pong ball. On the other hand, if two oppositely directed protons of equal energy are allowed to collide, then all of the energy of both protons is available in the reaction. In the third

[6] "Attainment of Very High Energy by Means of Intersecting Beams of Particles," by D. W. Kerst, et al. *Physical Review*, Vol. 102, April 15, 1956, p. 590.

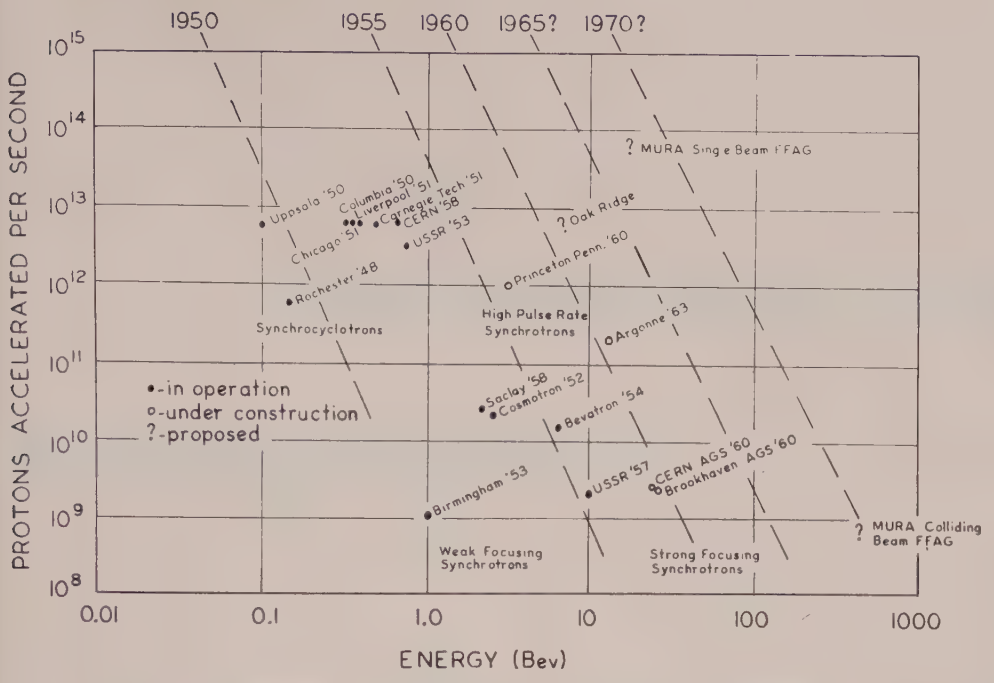


Fig. 11 — Beam intensity and energy of high-energy accelerators

column in the table we show the energy necessary in each of two colliding beams in order to achieve the same available energy. Thus, for example, if two beams of 15-bev protons were allowed to collide, the available energy would be 30 bev, an energy which could only be achieved with a conventional accelerator of 540 bev.

Several ways have been suggested of producing colliding beams of oppositely directed protons. One of the most ingeneous is due to the Japanese physicist, Ohkawa, who pointed out that it is possible to design a radial sector accelerator in which the positive and negative magnets are made exactly alike and in which oppositely circulating beams of protons can revolve simultaneously.[7] A model electron accelerator of this type which is now under construction at the MURA laboratory is shown in Fig. 10. This model will accelerate beams of electrons in opposite directions to energies of nearly 40 mev. It is expected that currents of 25–50 amp of oppositely circulating electrons will be achieved. If this can be done, it will be possible to study electron-electron collisions in which nearly 80 mev will be available. Without colliding beams, such collisions could only be produced with a 6-bev electron accelerator.

The graph shown in Fig. 11 summarizes the history and present status of the particle accelerator art. In this graph we plot horizontally the energy of an accelerator and vertically the number of protons which it is capable of accelerating per second. The solid dots represent accelerators already constructed. In each case the year of completion is indicated. The open circles represent the estimated performance of accelerators now under construction or contemplated together with the estimated dates of completion. In addition, we have indicated by question marks, the possible performance characteristics and possible completion dates for FFAG accelerators. The upper question mark represents a possible beam intensity from an FFAG accelerator at 15 bev. The question mark to the lower right represents the energy and beam intensity that would be required from a conventional accelerator to achieve collisions equivalent to those produced in a 15 bev colliding beam accelerator. The curves running down to the right through the graph indicate the maximum energies and beam intensities achieved as of various dates.

The development of particle accelerators has now reached the stage where very extensive development of theoretical and experimental techniques is required. As an example, the MURA laboratory utilizes an IBM 704 computer full time on accelerator problems. The spiral sector electron model shown in Fig. 8 was designed entirely with the aid of a digital computer. As a consequence, the dynamical properties of the electron orbits in this accelerator were well-understood even before the accelerator was constructed and the theoretical predictions have been amply confirmed by experiments with the accelerator. Experimental techniques in many areas, including high vacuum techniques, magnet design, radio frequency circuitry, and methods of particle detection, are required in order to build and utilize modern high-energy accelerators. To give just one example of the many technical problems which need to be overcome, we may consider that the alternating gradient accelerator now being constructed at Brookhaven has a diameter of 843 ft and that the magnets must be aligned around this circle within an accuracy of better than 0.020 in.

It is clear that high-energy particle accelerators represent a rapidly growing field which requires the best efforts of theoretical and experimental physicists and engineers.

[7] "Two-Beam Fixed Field Alternating Gradient Accelerator," by T. Ohkawa. *Review of Scientific Instruments*, Vol. 29, February, 1958, pp. 108–117.

some development problems with Large Cryogenic

Daniel A. Heald
Convair — Astronautics, General Dynamics Corp.

HEAT TRANSFER causes loading and starting design problems in large missile systems powered by cryogenic propellants. This manifests itself during loading as effective density variation, violent surface conditions, boiloff, and ice formation — problems which may be solved by insulating the tank. During starting it causes overheating and caviation — effects which may be reduced by recirculators and subcooled charge injections.

The study described in this paper centers around liquid oxygen and its variations in heat flux rate, which affect liquid density, surface condition, and replenishing requirements. The problem areas are made apparent by consideration of a hypothetical missile system.*

THE TITLE of this paper is sufficiently general as to require clarification of the term "Large Cryogenic Propellant Systems." Large should be taken to mean the size of the German V-2 and up. Cryogenic is used to mean very cold fluids such as ozone, methane, liquid oxygen, fluorine, and hydrogen which have boiling points of – 170 down to – 423 F. The discussion centers around liquid oxygen.

Associated with such systems are many challenging design areas such as structural problems arising from the use of materials at temperatures so low their ambient properties are grossly altered. But the region of heat-transfer problems are often beyond calculation; only extensive development testing can disclose the answers.

The areas which are discussed include variations in heat flux rate, which affect liquid density, surface condition, and replenishing requirements, which may result in insufficient propellant, engine operating problems, and changes pressurization requirements.

Propellant Loading Problems Arising from Heat Transfer

Consider a hypothetical missile system with liquid rocket engines using liquid oxygen as one of the propellants. The mission may include rapid loading and firing for military retaliation and boost functions on space projects. It appears rather straightforward when 0.5% accuracy on propellant weights at launch is requested or, stated another way, when it is necessary to fill the propellant tank to at least 98% capacity in order to assure maximum range. Let us suppose we are dealing with a missile system with a large vertical airborne propellant tank. The ground loading system, which must be considered an integral part of the whole, is a remote high-capacity system with extensive automatic controls. Why not just weigh the filled missile on load cells, which have an advertised accuracy of 0.1%? This approach may have to be abandoned when one considers that the gross weight of the launcher system may nearly equal that of the missile, that misalignment of the missile and forces from cross-winds introduce appreciable side loads which necessitate lateral load cells and complicate resolution of weighing system, that calibration of such a complex system is difficult, and accurate load cells are expensive and temperature sensitive, and that the propellant is rapidly boiling away.

* Paper presented at SAE National Aeronautic Meeting, New York, April 3, 1959.

Propellant Systems

Another approach is to calibrate the vessel, so that the volume is known for a given liquid height; then knowing the density and level, the propellant load can be calculated. This is the approach I would like to discuss, because it can lead into an interesting development program to establish these parameters.

We will discover that allowances must be made for the variables listed in Table 1. The reasons that such high percentage errors may be encountered during propellant loading will be discussed. Only careful system design and extensive development can lead to 0.5%, or maximum, propellant load on board at launch.

Effective Density — The easiest parameter to find is the relation between a given liquid height and missile tank volume. Accumulation of manufacturing tolerances may necessitate water calibration of each tank, which is possible to do with an accuracy of one-tenth per cent, using proper techniques and a precision positive displacement flowmeters or load cells. Then one should be able to calculate quite accurately tank shrinkage from the cold temperature of the propellant and tank expansion from internal pressure. A curve must be made up for each missile or liquid level versus volume.

Now let us consider an elusive variable: the density of liquid oxygen in a large missile tank. If one refers to available handbooks, LO_2 density is given at 71.2 lb per cu ft at – 297 F. If one checks carefully he finds that impurities are present in liquid oxygen, such as nitrogen which may be left over from manufacturing or introduced during storage and transfer cycles if nitrogen gas was used for pressurization. This may alter the density appreciably in the order of 0.1% or more, and is difficult to measure accurately in field service.

Referring to Fig. 1, it is apparent that liquid oxygen density varies markedly with temperature. For example, if the temperature has increased 8 F from STP boiling (8 deg warmer than – 297 F, or 8 deg ΔT) the density has dropped over 2%. Conversely, if the liquid oxygen is at an altitude above sea level — say at Edwards Air Force Base which is about 2500 ft — then the ambient boiling temperature is nearly – 300 F, so the density is about 0.5% heavier. From this it is clear that the exact temperature of the propellant must be known in order to calculate the total mass accurately. But we shall see that the heat-transfer rates and temperatures vary greatly in a large tank.

Why is the average or "effective density" different from that of a chem lab sample?

In large systems where gravitational forces are

Table 1 — Variables in Propellant Load Calculations

Propellant Load Parameters	Probable Percentage of Error
1. Missile tank section area/calibration	0.1 – 1.0
2. Liquid level indication	0.1 – 3.0
3. Propellant density, at ambient boiling	0 – 0.5
4. Propellant density, effective	0.2 – 2.0
5. Topping/replenishing system and controls	0.1 – 0.5
6. Ice formation	0 – 1.0
Total	0.5 – 8.0%

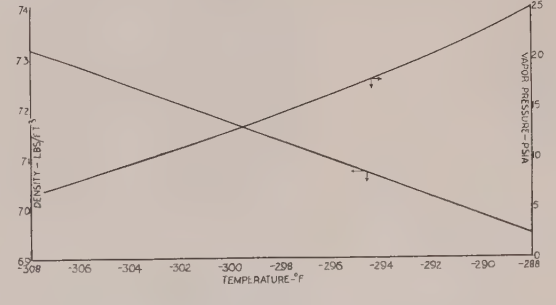


Fig. 1 — Liquid oxygen density and vapor pressure

appreciable, there is an increase in pressure in the liquid proportional to the liquid head. Without circulation, the liquid temperature will increase as a function of distance below the surface until at any point the vapor pressure is equal to local pressure. But convection tends to mix the liquid. Elementary experiments on bare tanks indicate no measurable difference in temperature throughout the tank. Heating along the walls and bottom bulkhead of the tank will generate a film of warm liquid (and perhaps some bubbles) which lowers the local density and perhaps permeates the bulk.

When appreciable lift affect is caused by bubbles and rising convection currents, the surface level will be raised. Pressurization may collapse the bubbles, so the surface settles noticeably, and tank volume is wasted.

Fig. 2 shows the warmer outer fluid rising to the top, boiling, and thus reducing the temperature of the surface liquid, which then settles and mixes. In a vented tank the top several feet will be saturated liquid, but most of the remaining propellant may be several degrees warmer. Under adverse environment, this affect may become markedly increased. This is such an unstable condition that temperature profiles are very difficult to obtain and the exact mechanics of the convection can not be described mathematically with any assurance. Extensive testing with the actual tank and propellant will probably disclose an effective density 0.5% lower than expected even in average weather.

Surface Conditions — This recirculation has disadvantages: violent surface activity in a vented tank. If one wishes to use a simple loading control, such as a probe to indicate surface level, then he must consider the conditions at the surface. Even under mild atmospheric conditions the liquid may be froathing and splashing several inches. As atmospheric conditions become more severe — wind or rain, the surface may become so violent as to slosh several feet up the tank walls and out the vent valve, even though the tank is only 90% full. It is usually considered unsafe practice to spray 100 gpm or so of cryogenic fluid out the vent valve onto the missile exterior or launch complex.

This may force missile system complication by abandoning a simple level indicating device and turning to a very precise ΔP measuring device. Baffles may have to be added across the vent port, or an overflow duct added to the ground support equipment.

To study surface conditions, it is advantageous to get photographic coverage. At best the results are qualitative, but one can gain considerable insight into the entire heat-transfer process and a healthy respect for the problems involved. With movies one can see that the convection currents move at several feet per second and surface agitation can be violent.

Fig. 3 shows approximate effects of various environments encountered prior to launch on the heat transfer through a cryogenic propellant tank wall. A ratio of heat flux equal to one is taken as the rate into a tank with average insulation such as spun glass matting (not a vacuum). The curves are approximate ratios for unprotected tanks. Curve A is for zero wind, indicating that frost formed from atmospheric moisture creates a fair insulation. The higher the partial pressure of water vapor, the heavier the frost, the less the heat flux. Curve B is for rain conditions, where the heat transfer is high until ice formation reduces the rate. The higher the rainfall and/or wind velocity, the higher the initial rate. Curve C indicates that a very high wind, such as 60 knots with no rain, causes heat flux rates as much as 40 times as high as that for an insulated tank.

Topping System — The level — wherever it is — is bound to be dropping due to boiloff. Rates which may be about 1% per hour under average atmospheric conditions may increase many fold in a 60-knot wind. Realize that the nominal ullage space is small to avoid wasted tankage on a high-performance vehicle — possibly only 1% of total tank volume — and the replenishing system is ground support equipment remote from the missile.

For the average conditions, a low replenishing rate is required, so the piping must be of small diameter and well insulated to avoid liquid overheating before it reaches the missile. But under adverse weather, a much higher capacity, more rapid response system is required. Precise level or

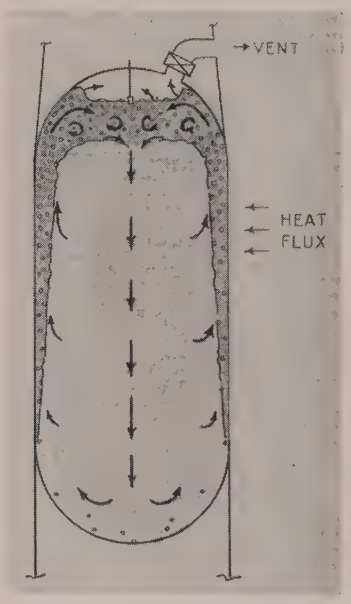


Fig. 2 — Convection paths

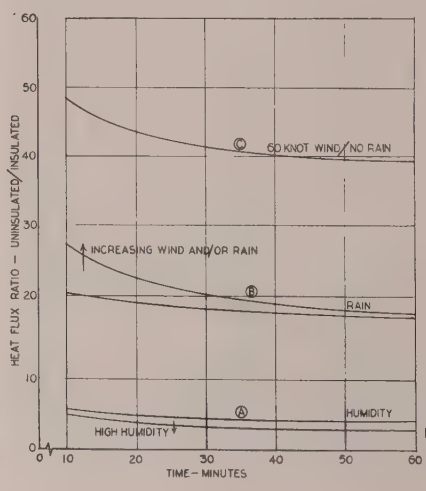


Fig. 3 — Heat flux into unprotected cryogenic propellant tank for various environments

mass measuring equipment is required, which is a complex task as indicated before. And a modulating valve of unusually fast response with unusually wide throttling range may be necessary. Only careful integration of the missileborne and ground support equipment, supplemented by extensive testing can produce such a system.

Ice — Cryogenic propellants add one further complication: Will the ice formation on the outside of an uninsulated liquid oxygen tank be heavy enough to affect trajectories? We are accustomed to seeing a powdery frost on the exterior of liquid oxygen tanks, but this is very light and tends to fall away as snow during launch. A light rain produces ice formations, mostly on the windward side with thicknesses up to 0.25 in. But tests run to simulate heavy rainfall of several inches per hour can produce a layer of ice over 2 in. thick. If such a coat were retained until melted by aerodynamic heating, a great weight penalty is being paid. Will vibrations and tank expansion due to pressurization crack off the ice? If so, will falling ice chunks do any damage? Can a thin coating be applied to the missile which will discourage ice retention? While these questions are being investigated, a new approach may be worthwhile.

Environmental Protection — One then begins to consider insulating the tank. For liquid oxygen it would appear best to release the environmental protection at launch, but for the liquid hydrogen it may not be advisable to jettison the protection until high altitude because of high surface liquid weight ratio. In either case one must be assured that the insulation will release—that it will not freeze to the tank. The German V-2 had permanent spun glass insulation between the pressure vessel and the outer shell, but this is heavy! Completely shrouding the Martin Viking Oxidizer Tank was a blanket which was jettisoned just prior to launch. Other means, such as enveloping the entire missile and launcher in a thin mylar balloon may be considered, but this also becomes complicated when high winds are considered. If insulation is necessary to prevent ice formation, or to reduce aerodynamic heating, then will it help solve some of the problems previously discussed?

Insulation can yield the following benefits:

1. Greatly reduced boiloff with resulting simplification in replenishing system and controls. Boiloff is virtually independent of environment.
2. Nearly constant effective density, that is, less dependent on atmospheric conditions.
3. Less violent surface conditions making level indication less complex.
4. Virtual elimination of ice formation.

But the price may be high: The blanket and jettison system may be a major development program in itself, and still evolve complex and undependable. Will it freeze to the tank? What release mechanism is reliable? It may be an operational handicap.

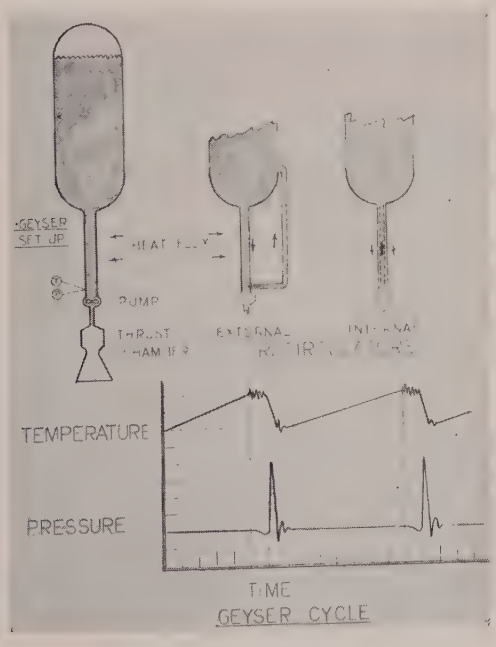


Fig. 4 — Overheating

Engine Starting Problems Caused by Heat Transfer

Feed Line Overheating — In an arrangement where a long feed line connects the cryogenic propellant to the engine system, more heat-transfer problems should be expected, such as overheating, geysering, or voiding.

The line has a higher surface-to-volume ratio than the main tank, so temperature rise will be more rapid under the same atmospheric conditions. No circulation is present as there is in larger tanks. In a 60-knot wind, temperature may rise 4 F/min in an unprotected LO_2 line. Unless a fix is developed, the overheating may lead to geysering which is similar to perculating in a coffee pot. The duct can be partially voided of liquid due to explosive boiling, ordinarily triggered near the top of the line and progressing rapidly to the bottom. The system will refill with liquid once boiling has subsided, causing a pressure surge many times the initial pressure. These transients have been known to deform or rupture missile plumbing. On the other hand, experience has shown that large portions of a ducting system can boil continuously, without explosive geysering.

The usual remedy is to provide bleeds, but this costs weight. In a large system the tank might have to be several hundred gallons oversize to accommodate the bleed. This usually consumes time at a critical moment in the starting sequence and may require additional instrumentation.

Development testing can evolve several types of solutions, none of which can be conclusively calculated. A recirculator usually proves advantageous. The external recirculator may also be a weight penalty unless there is more than one system to feed and cross-over plumbing is planned anyway. There may have to be some difference in heat-transfer rates in the two lines to promote circulation — either materials with different conductivities, one insulated or otherwise protected from the atmosphere, or different line sizes.

Or an internal recirculator may be used where the inner pipe carries cold liquid down while the outer ring is subject to heating and rises.

In either case it may help to induce circulation by injecting gas (airlift pump effect). Fig. 4 shows a

typical ducting arrangement which may be subject to overheating and geysering. A pressure/temperature history during a geyser is shown. Notice that the temperature slowly rises to saturation point, and boiling commences. The boiling becomes so violent, there is an eruption, forcing the liquid out of the duct. Overshoot causes slight decrease in pressure and as the gas escapes, liquid falls back into the duct creating a sharp pressure spike. The figure also indicates types of recirculators which may eliminate geysering.

Engine Pump NPSH — Let us now consider the engine pump net positive head, which is also a function of propellant temperature. Remembering that suction head above vapor pressure, in feet equals

$$NPSH = H_{liquid} \times G + H_{vel} + H - H \text{ friction losses} = H \text{ vapor pressure, ullage pressure}$$

close attention must be paid to the vapor pressure in cryogenic systems. It can become the key to starting problems both on the ground and starting or re-starting in flight.

Visualize a ground start situation, where pump buildup occurs in less than a second. In order to avoid severe cavitation, the available head just prior to start may have to be triple the steady running requirement to compensate for acceleration losses. The designer might increase ullage pressures and/or grossly enlarge plumbing to reduce inertia losses, both with resultant high increase in airborne weight.

However, it may sometimes be possible to achieve the same results without penalty to airborne weight by decreasing vapor pressure at least locally.

Subcooled Charge — This can be done by injection, near the pump inlets, of subcooled propellant. Liquid oxygen chilled about 20 F below its ambient boiling temperature has a 10 psi lower vapor pressure with little or no loss in airborne weight. The subcooling may be done by evacuating the ullage above the storage vessel, by storing the liquid oxygen in a liquid nitrogen shroud or by passing it through a liquid nitrogen heat exchanger.

A liquid nitrogen/liquid oxygen heat exchanger can be a surprisingly simple device. By boiling 1 lb of LN_2, about 10 lb of LO_2 can be chilled from – 293 to – 315 F, even at topping flowrates.

Local Cavitation — Cavitation must always be considered in any system flowing a liquid near its saturation point. In complex manifold systems containing protuberances in the line, such as throttle valves, flow vanes, or turbine-type flowmeters, local cavitation can occur even when the average line pressure is well above vapor pressure. For example, if the apparent line pressure is 50 psia, but the vapor pressure of the liquid is 25 psi — which is not unusual in a liquid oxygen system — turbulence may be so severe after a valve, so that local boiling occurs. Will these vapor bubbles collapse before reaching the engine? If not, will they adversely affect pump performance, or even choke the feed line? Conventional instrumentation will frequently not yield the answers, so more sophistocated apparatus may be required for cavitation analysis.

Development Test Program

When it becomes necessary to determine accurately temperature distribution, effective density, overheating tendencies, and boiling rates discussed above, a development program must be planned. To get a general feel for the various phenomena without excessive costs, a model program using liquid Freon may be valuable. A small scale model of the missile propellant tank, feed lines, and dummy engine pumps may be constructed of transparent material such as pyrex, Vycor, or plexiglass. Realistic filling and pressurization provisions should be made. Various environments can be simulated with heat lamps, fans, and water sprinklers. But the final answers can only be accurately obtained from the full-scale prepellant system using the cryogenic liquid involved. The cost of fans and engines to create a 60-knot wind may be prohibitive, so the effects of this adverse environment may have to be extrapolated from data taken during a natural storm, which may require a long wait in "Sunny Southern California." It is possible to photograph the liquid surface or bulk activity with camera and light ports carefully purged to avoid frosting/freezing. Precise ΔP manometers, a string of vapor/liquid detectors, a mass of temperature transducers, and cavitation detectors are required for anything more than qualitative results. Even after such a program, some questions such as ice retention may have to be investigated during missile firings.

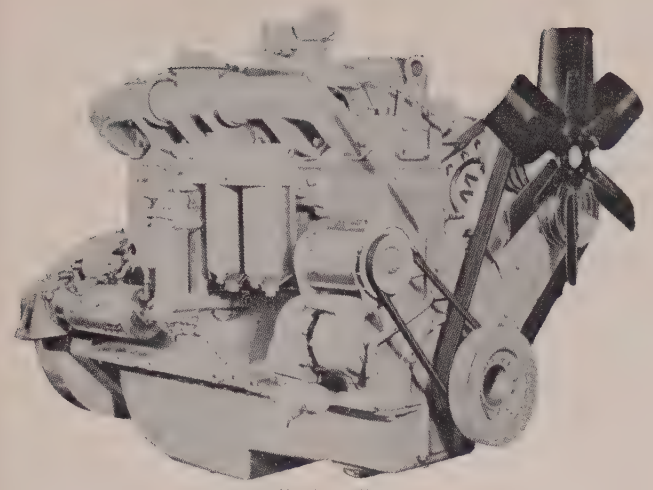
Flight Problems

Self-Pressurization — If a cryogenic propellant is loaded into the missile at nearly ambient boiling temperature, then pressurized and held for even a few minutes in adverse weather, the vapor pressure may markedly increase. Aerodynamic heating during booster flight may also cause rapid bulk temperature rise. Or long periods of solar radiation may have the same effect. It may reach the saturation point.

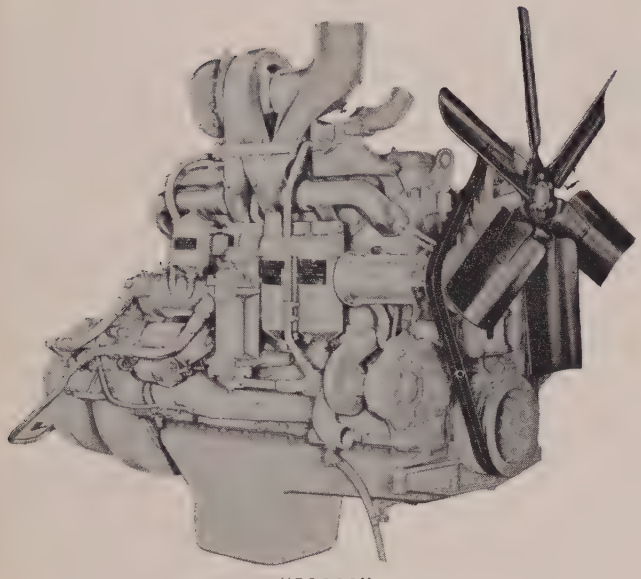
One may take advantage of this fact to allow the system to be self-pressurizing, thus possibly saving weight of auxiliary pressurization systems. However, this gain may be offset by engine pump starting problems if the propellant is a saturated liquid and no external acceleration is applied.

Zero G Affects — Let us consider a space vehicle which must have in-flight engine starting capabilities, perhaps to decelerate onto a planet, after days of coasting. If the tank had been half full of cryogenic propellant, what is its condition just prior to restart? After several days in a zero G field, subject to solar radiation, the liquid will be saturated and probably dispersed everywhere in the tank except at the pump inlet. It may be a fog rather than pure liquid. Will positive expulsion pistons in the main tank be necessary? Will it be necessary to start the engine from small start tanks to gain acceleration to settle the main propellants? Must auxiliary propulsion be provided to settle the propellants? If so, how many G's will be required? How will the feed lines, which are doubtless void and hot, be refilled so the pumps can prime? How can one test for such zero G conditions on the ground?

These questions must be answered by revolutionary types of development testing if cryogenic liquid propellants are to hold their place in future rocketry.



"16000"



"21000"

Fig. 1 — Direct-injection diesel engines

development of two new Allis-Chalmers Diesel Engines

Hans L. Wittek

Allis-Chalmers Manufacturing Co.

THIS PAPER deals with the design and development of the naturally aspirated and turbocharged versions of a direct-injection diesel engine, known as the Allis-Chalmers 16000 and 21000 (Fig. 1).

Because these engines were intended specifically to power certain crawler and wheel-type vehicles built by our company, the original design and performance specifications were defined by the space available in these machines and by the power and torque characteristics required by them. In addition, the engines had to be designed and arranged so that they would be suitable not only for a variety of commercial applications such as generator sets, shovels, compressors, and the like, but also for oil field and marine service.

The information on desired performance, supplied to us at the beginning, is shown in Table 1. Furthermore, it was stipulated that the new engines had to have the characteristics shown in Table 2.

The new direct-injection diesel engines which resulted from this program were officially announced in December, 1958, but had, by then, been in actual customer service for more than six months. We feel that sufficient operating data is now available to permit a fair evaluation of what has been achieved. At this time, therefore, it seems appropriate to discuss the thinking on which the design was based, the various stages of development, and the results obtained.

Basic Design Considerations

It is not the intention of this paper to make it appear that our solution is the best or the only one possible. We would be most reluctant to claim that

THE NEW DIRECT-INJECTION diesel engines — the naturally aspirated 16000 and the turbocharged 21000 — were developed to power specific crawler and wheel vehicles built by Allis-Chalmers. Thus the original design and performance specifications were defined by the space available in these machines, and by the power and torque characteristics required by them. Also, the engines had to be suitable not only for commercial applications such as generator sets, shovels, and compressors, but also for oil field and marine service.

Torque and speed requirements indicated that a 5¼-in. bore and 6½-in. stroke would give the desired performance. To meet the low heat rejection and good starting requirements, an open-chamber combustion system had to be used. The three-valve arrangement — two intake and one exhaust — was chosen because it offered low pumping losses and reasonable cost.

This paper describes the design considerations and development work which produced the new diesel engines.*

perfection has been reached — there are, after all, quite a few design variables each independently alterable, and every one having a pronounced influence upon performance. If there were just 10 variables and if five variants of each of these had to be tested with each variant of every other variable, and if we could complete four tests during every working hour, then, to find the one perfect combination would take just about 187 years.

To start at the beginning of this development, we must go back about eight years.

Many thousands of our naturally aspirated, supercharged, and turbocharged engines (known as the DA, DAS, and DAT-844 models, respectively) have been built and are in service. These energy-cell-type engines have some very desirable characteristics, particularly smoothness of operation, good flexibility, and extreme reliability. However, research work aimed at a substantial increase in output combined with improved fuel economy, showed that these demands could not be satisfied by this basic engine type. Consequently, an entirely new development had to be undertaken. It was a foregone conclusion that the new engines, in addition to fulfilling the requirements mentioned previously, also would have to retain the good qualities of the older models.

Before deciding on any definite basic approach, several design studies were made and the most modern diesel engines of similar speed and size built here and abroad were carefully investigated. After this had been done, actual work was started early in 1952 when a definite item, covering the development of a direct-injection combustion system, appeared in our engineering budget for the first time.

Torque and speed requirements having been established, the mean effective pressures, suitable for each specific application and obtainable from an as yet undefined combustion system, had to be predicted. Available machine tool equipment made it desirable to adhere to existing cylinder dimensions, but we had specific instructions from management not to let this consideration influence our thinking and, should it prove necessary, to request entirely new manufacturing facilities.

After lengthy calculations and estimates, we felt confident that a 5¼-in. bore and 6½-in. stroke would give the desired performance. Analysis showed that a crankshaft of sufficiently greater bending strength could be accommodated within the existing cylinder center distance. Consequently, it was decided that cylinder block tooling could be retained and the design was begun on this basis.

The question of what basic type of combustion chamber to employ was approached with entirely open minds. However, it was obvious that good starting and low rate of heat rejection could be obtained only through the use of an open-chamber combustion system. As mentioned, we had tested quite a few different types of direct-injection engines. The performance of one of these, built by A. E. C. Limited of England, was so outstanding that it induced us to pay serious attention to the general type of combustion chamber used by them. We will discuss later the considerations which made us decide exactly what course to take.

Management had, as already mentioned, specified four basic requirements. We in Engineering added a few of our own, so that we now had a list which looked about as shown in Table 3.

It will be noted that "low specific fuel consumption" is not shown as a separate item in this specification. It was not necessary to mention it as smoke-free operation and low heat rejection can be achieved only if a substantial part of the fuel admitted into the engine is converted into useful work. Thus, high economy is implicit in the characteristics asked for.

It is a well-known fact that hardly any piece of an engine can be altered without setting off a chain reaction of other necessary changes. This is particularly true of anything connected with the combustion system. Each requirement mentioned not only needed its own special treatment but it also influenced most of the others.

Before we could start on any actual design work, we had to decide on the number of valves to be used. Previous experimental work on an energy-cell-type machine had shown that, in an engine of this size, the 3-valve arrangement (two intake and one exhaust) results in a happy combination

* Paper presented at SAE Summer Meeting, Atlantic City, June 16, 1959.

of low pumping losses and reasonable cost. It was agreed, therefore, that this was to be the arrangement.

Combustion Chamber

Having been accustomed to the very large single hole of the pintle nozzle, we arbitrarily established that the new nozzle holes had to have a diameter of at least 0.012-in., as we felt that this was about as small as could safely be entrusted to the average mechanic.

The anticipated fuel quantity per cycle, together with the established hole area and the desirable duration of injection of approximately 25-deg crank angle, showed that four or five nozzle holes were required. The lower of these numbers was chosen so as to result in the largest possible nozzle hole diameter.

The fundamental problem in the design of any combustion chamber having internal mixture formation simply consists in offering to the fuel, at the right moment, all the air which is intended for it. This means that a particle of air located near the periphery of a circular combustion chamber must cover, while injection lasts, a distance equal to the circumference of the chamber divided by the number of sprays (Fig. 2).

This can be expressed:

$$\frac{2r_c\pi}{12h} = v_{tc} \times t \qquad (1)$$

where:

r_c = Combustion-chamber radius, in.
h = Number of spray holes
v_{tc} = Tangential air velocity in the combustion chamber, fps
t = Duration of injection, sec

If nothing is done to increase the tangential air velocity in the combustion chamber after the air has been admitted to the cylinder, then the tangential admission velocity (v_{tv}) has to be at least equal to the required swirl, provided that the mean distance of the entering airstream from the cylinder axis is equal to the radius of the combusion chamber. If this is not so, it could be assumed that the condition:

$$v_{tc} \times r_c = v_{tv} \times r_v \qquad (2)$$

would have to be fulfilled so as to produce the needed swirl.

t can be expressed as a function of engine speed and angular duration of injection a. If N is rpm:

$$t = \frac{60}{N} \times \frac{a}{360} \qquad (3)$$

Thus, there results:

$$\frac{2\,r_c\pi}{12\,h} = v_{tc} \times \frac{60}{N} \times \frac{a}{360} \qquad (4)$$

and substituting:

$$v_{tc} = v_{tc} \times \frac{r_v}{r_c} \qquad (5)$$

$$\frac{r_c\pi}{h} = v_{tc} \times \frac{r_v}{r_c} \times \frac{a}{N} \qquad (6)$$

$$v_{tv} = \frac{r_c^2}{h} \times \frac{\pi N}{r_v a} \qquad (7)$$

If the combustion chamber were a flat disc, having the cylinder bore for its diameter, then for an engine having 25-deg duration of injection, 5.25-in. bore, four injector holes, an inlet valve located 1.75 in. from the axis of the cylinder, and operating at 2000 rpm, the theoretically necessary tangential inlet velocity would be:

$$v_{tv} = \frac{5.25^2}{2^2 \times 4} \times \frac{\pi 2000}{1.75 \times 25} = 248 \text{ fps}$$

This result would be slightly modified by taking into consideration the spray dispersal angle. Assuming it to be 10 deg, v_{tv} would become:

$$248 \times \frac{80}{90} = 220 \text{ fps}$$

To fill the cylinder, the inlet airstream must obviously have a vertical component in addition to the tangential one, in other words it must slant downwards into the cylinder. If it formed a 45-deg angle with the cylinder axis, the minimum air velocity at the point of entry into the cylinder would then have to be $220 \times \sqrt{2}$, equal to some 312 fps.

All this was based on minimum required tangen-

Table 1 — Performance Requirements

Application	Type	Rated Power		Peak Torque	
		bhp	rpm	lb ft	rpm
HD-16	Crawler	155	1600	552	1200
HD-16C	Crawler	165	1800	524	1200
TS-260	Wheel	220	2000	643	1500
HD-21	Crawler	240	1825	702	1600

Table 2 — General Requirements

1. Stamina as needed in construction machinery service.
2. Good starting in low temperatures.
3. Smoke-free operation.
4. Low heat rejection.

Table 3 — Final Requirements

1. Stamina needed in construction service.
2. Good starting in low temperatures.
3. Smoke-free operation.
4. Low heat rejection.
5. Potentially high specific output.
6. Simple, plain and symmetrical cylinder head.
7. Good flexibility and wide speed range without avoidable complications.
8. Relatively large injector holes.

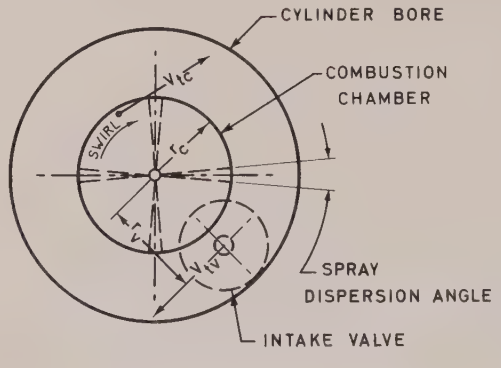


Fig. 2 — Swirl

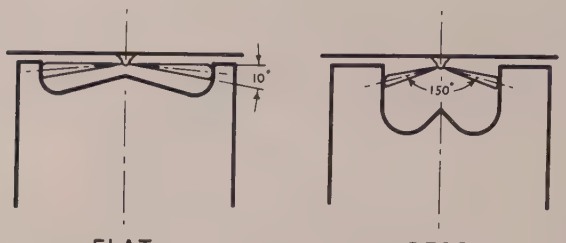


Fig. 3 — Combustion chambers

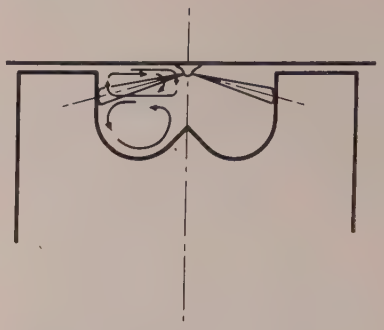


Fig. 6 — Secondary air movement

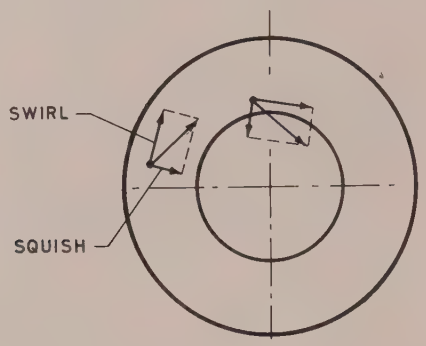


Fig. 4 — Swirl and squish

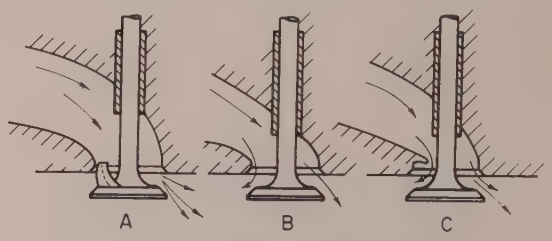


Fig. 7 — Methods of producing swirl

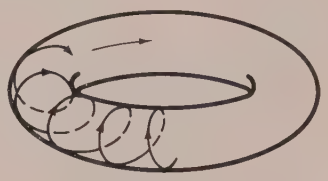


Fig. 5 — Rolling torus

tial swirl velocity at the time of injection, but it is reasonable to assume that the air will be slowed down substantially during the relatively long time which it spends in the cylinder prior to injection, that is, during the whole of the compression stroke. It is anybody's guess how much this slowing down would amount to; but, intentionally underestimating it at about one-third of the initial value, we see that in the example cited the necessary mean inlet air velocity would work out to be near 465 fps.

It is realized that the foregoing calculation is not only oversimplified but also that the assumption of the existence of a forced vortex instead of a free one, which is equally likely to exist, is open to contention. We must also remember that the path of a fuel droplet cannot be a straight radial line. It is deflected by the air movement and, as soon as the droplet acquires a rotary motion around the cylinder axis, it becomes subject to centrifugal forces. This is the reason why experimental data on spray penetration, obtained in a quiescent atmosphere, are not of much help if a rapidly rotating vortex exists.

In this connection it should be mentioned that modern data on spray penetration is sadly lacking. Apparently very little has been done to extend P. Schweitzer's outstanding work of 1937.[1] The information which he supplied is, of course, still valid, but as his research was performed to fill the needs of that time, it did not extend to the high spray velocities and chamber densities used today.

In spite of the above reservations, the statement made by the equation, namely that required swirl velocity is proportional to the square of the combustion-chamber diameter and inversely proportional to the number of spray holes, agrees quite well with experimental evidence.

An actual air velocity anywhere near the one mentioned would, of course, have a detrimental effect upon volumetric efficiency, so that we have to find means either to reduce the required swirl velocity or to increase the swirl after the air has passed through the intake valve, or both.

We have seen that, for a given number of nozzle holes, engine speed, and combustion-chamber diameter, the necessary swirl velocity increases as the square of the latter parameter. The obvious answer, then, is to reduce this diameter.

The benefit derived from doing this is twofold: the required tangential speed is reduced and, thanks to the law of conservation of energy, the angular velocity of the swirl is increased.

Theoretically, this increase should be inversely proportional to the ratio of squares of the two diameters so that the swirl should increase four times if the diameter of the rotating body of air is halved. In reality, the gain in swirl is not that great, being somewhere near the direct ratio of radii, but it is, nevertheless, very worthwhile. It means that, for any required swirl rate, the needed air velocity at the valve can be reduced by a little more than one-half, if the combustion-chamber diameter is about equal to one-half the bore.

As we assume that the compression ratio is fixed, a decrease in the diameter of the combusion chamber requires an increase in its depth. As shown in Fig. 3, the shape of the chamber no longer conforms to that of the spray. If only the primary air move-

[1] "Mechanism of Disintegration of Liquid Jets," by P. H. Schweitzer. *Journal of Applied Physics*, Vol. 8, 1937, p. 513.
"Penetration of Oil Sprays in Dense Air," by P. H. Schweitzer. Pennsylvania State College Bulletin No. 20, 1934; and No. 46, 1937.

ment, the swirl, existed so that the air content of the chamber rotated like a solid disk, we would not achieve our basic objective, namely to pass all air through the fuel sprays while injection takes place. However, the smaller diameter of the chamber, or rather the remaining flat part of the piston crown, quite automatically supplies the needed secondary air movement, the squish.

Bruno Loeffler[2] has developed a very neat and simple equation which shows that theoretical squish velocity is proportional to the difference between the squares of cylinder bore and cavity diameter. Thus, conservation of momentum gives us the tangential component and the Loeffler equation the radial component of the airflow as it enters the cavity (Fig. 4). Experimental evidence shows that the air actually knows what it is supposed to do.

One could now reach the conclusion that the air contained in the cavity forms a torus which rotates around its own axis and which simultaneously rolls around the center of its circular cross-section (Fig 5). Actually, the air movement is much more complicated, largely due to the effect of air already in the cavity and rotating at high speed, upon other air just entering and which has not as yet been accelerated. In addition, the air, due to its swirl, is subject to centrifugal forces and to friction where it contacts the wall.

If someone were rotating with the induced swirl, the air movement would look to him as shown in Fig. 6.

It can be seen that, depending on its distance from the nozzle, a fuel droplet is contacted by air travelling across, as well as with, the direction of its own movement. Obviously, then, whatever are the conditions conducive to short ignition delay, they will be fulfilled somewhere along the line. Proximity to a wall and the temperature of the latter can also contribute to shortening the delay period. Some experiments of ours indicated an unexpectedly short delay. We thoroughly distrusted our observations until we read the findings recently published by A. Pischinger and F. Pischinger.[3]

Everything which has been said above would apply equally well to a spheroidal or ellipsoidal combustion chamber, but manufacturing a cavity of such a shape and controlling its exact size and finish is not an easy matter. We felt that a cylindrical cavity, having straight walls, should perform equally well. The performance obtained from our combustion chamber, which we believe to be at least equal to anything that has been claimed for one of the more complicated forms, has vindicated our opinion. A possible argument against a chamber of small diameter might be based on the probability that some fuel, still in liquid form, might reach the wall. Not very long ago fuel impingement was looked upon as something which had to be avoided at all cost. Meurer's[4] findings, however, induced all of us to take another look.

It is a matter of course that fuel impingement cannot be used indiscriminately. The fuel must not be permitted to reach the cylinder wall. Angle of impingement upon the wall of the cavity as well as wall temperature at the point of impingement must be exactly correct so as to prevent smoke and objectionable smell. The deep combustion chamber of cylindrical form provides quite a target for any desired impingement without the risk of liquid fuel reaching the cylinder wall. The gradient in wall temperature, highest at the top and lowest at the bottom of the cavity, enables one to choose the correct wall temperature at the place of impingement by selecting the spray cone angle. If this angle is also effective so far as the existing airflow pattern is concerned, the arrangement is a most effective one.

J. Bottger[5] and A. Jante[6] have just published some findings which show that considerable work is being done on partial fuel impingement, as distinguished from the complete impingement use in the "M" engine. It would be outside the scope of this paper to report on our research aimed in this direction. However, our efforts to explain just what is taking place in our particular type of combustion chamber are progressing, so that we will shortly be able to speak about its combustion kinetics and its chemistry of combustion.

We have seen that the effort to combine a small number of large nozzle holes with low inlet velocity lead us to a compact chamber design. This, in turn, benefits starting ability. The cold cylinder head and piston surfaces are much farther away from the fuel spray than if the shape of these walls is made to conform to that of the spray.

The compact shape of the chamber, which is something very different from surface-to-volume ratio, permits the formation of a hot core. That this is true is shown by the beneficial influence of the "Pintaux" nozzle upon the starting ability of the Ricardo[7] "Comet III." Instantaneous temperatures are difficult to measure, particularly when conditions are as unstable as during the cranking of a cold engine, but we know that the temperature of the core is at least 160 F above the mean compression temperature when the wall temperature is around 30 F and the cranking speed 150 rpm. What is equally important is that the hottest point is located about one-half of the cavity radius outwards from its axis and about one-third of its depth above the bottom. Thus, the lower edge of the spray passes through the area having the highest temperature.

Initial Swirl

It remained to be decided just how the required initial swirl was to be produced. The various possible methods are well-known. (See Fig. 7.) Streamlined or directional ports (B), inserts or protrusions (C) within the port just above the valves, or masked valves (A) all give the desired result. Of these the directional port appears to be the simplest and most straight-forward solution. Actually, this is not quite so as we learned that with this arrangement we could not produce entirely satisfactory behavior over the whole speed range without resorting

---

[2] "Development of Improved Automotive Diesel Combustion System," by Bruno Loeffler. SAE Transactions, Vol. 62, 1954, pp. 243–264.

[3] "Bombenversuche uber die Diesel — Verbrennung unter motorischen Bedingungen," and "Der Einfluss der Wand bei der Verbrennung eines Brenstoffstrahles in einem Luftwirbel," by A. Pischinger and F. Pischinger. MTZ, Vol. 20, January, 1959.

[4] "Evaluation of Reaction Kinetics Eliminates Diesel Knock — The M-Combustion System of MAN," by J. S. Meurer. SAE Transactions, Vol. 64, 1956, pp. 250–272.

[5] Article by J. Bottger. *Maschinenbautechnik,* Vol. 7, September, 1958, p. 482.

[6] Article by A. Jante. *Kraftfahzeugtechinik,* Vol. 8, August, 1958, p. 284.

[7] "High-Speed Internal-Combustion Engine," by Sir Harry R. Ricardo. Pub. by Blackie & Son, Ltd., London, 1953.

to variable injection timing. The same held good for guiding members inserted or cast into the intake port. Without claiming that it provides a completely satisfactory explanation, this illustration nevertheless points up a probable reason. The valve mask retains its effectiveness during the entire intake cycle, whereas the other methods are bound to become less and less effective the farther the valve is off its seat. They are least effective at the time when control is most needed, that is, at full valve lift and highest mass flow.

No matter what arrangement is used, the desired swirl can be quite easily produced for one specific engine speed, but retaining satisfactory conditions over the whole speed range is much more difficult. When no mask is employed, directional control becomes the less effective the lower the engine speed because lower velocity through the throat will permit some air to enter in an undesirable direction. Thus, it will counteract rather than augment the desired swirl.

This can be compensated for to some extent by retarding injection. This means that if proper performance over a wide speed range is desired, a choice has to be made between the added complication and high cost of a variable injection timing device or the slightly increased cost of a masked intake valve. The higher valve cost just mentioned is the only drawback of the masked valve. Many thousands of engines equipped with such intake valves (engines built not only by us but by A. E. C., Saurer, M. A. N., Fiat, and the rest) furnish proof that the life of the non-rotating intake valve is equal to that of the conventional type.

Some of the foregoing should be qualified insofar as there seems to exist a definite size influence. In engines of smaller bores, say below 5 in. or so, satisfactory swirl conditions over a fairly wide speed range are easier to obtain without a masked valve than when the bore becomes larger. This is particularly true when only a single intake valve per port is used.

It is often assumed that the mask has an undesirable effect upon volumetric efficiency which the other methods lack. Referring again to the diagram one sees that this is not the case whenever a directional flow is to be obtained. Such directed flow through a conventional poppet valve can, obviously, be produced only if part of the free valve cone does not carry any flow so that all air is forced to pass only through a part of the periphery of the cone. It stands to reason that, the closer to the final orifice, that is, to the valve head, the means for imposing direction are located, the smaller a restriction will be needed to obtain the desired result.

The actual magnitude of the effect of the mask upon flow is surprisingly small as has been shown by C. B. Dicksee.[8]

Valve Timing

Having agreed on the basic shape of the combustion chamber, namely on the use of a straight-sided and relatively deep piston cavity, the approximate valve timing, that is either a smaller or a large valve overlap, had to be decided upon. Since we were aiming at fairly high mean effective pressures, the standard answer would have been in favor of large overlap.

We have all been taught that, in the case of supercharged engines, a large overlap results in all kinds of benefits such as better scavenging, internal cooling, and the like. We did wonder, however, if the benefits derived from considerable overlap might not be outweighed by their disadvantages. It seemed possible that only turbine inlet temperature benefits from this arrangement to a noticeable degree, but that it might not necessarily produce a worthwhile reduction in piston and valve temperatures. Assuming then that, with little overlap, the valve temperature and turbine inlet temperature at the highest mep desired stayed well below acceptable limits, could not any needed reduction in piston temperature be more easily obtained by liquid cooling? If so, would not the latter method obtain the desired goal from a smaller expenditure of energy.

We were somewhat reluctant to assume from the start that a substantial valve overlap was a must, simply for the reason that we had had experience with valve clearance recesses in the piston and, therefore, wished to avoid them if possible.

It can hardly be denied that valve clearance recesses, by destroying the symmetry of the piston crown, increase the likelihood of thermal stress concentrations. The exposed edges add to the surface and the variation in piston wall height above the top ring tends to increase temperature variations in the upper ring belt. The latter condition can, of course, be corrected by increasing the piston height above the pin (which increases engine weight), or by more effective internal cooling of the piston.

There is no question that the design problems inherent in piston recesses can be solved, as is proved by the fact that some extremely successful engines have this feature. What finally decided us to attempt a valve timing which does not require recesses (Fig. 8) was the consideration that, the more cut-up the piston crown, the more initial swirl is needed to obtain the required final swirl rate. If more initial swirl has to be imposed on the incoming air, more energy is required to achieve this result which means that pumping losses increase and volumetric efficiency decreases.

So far, we have no reason to regret our decision, as there is no indication of any need for more overlap up to the highest mean pressures so far obtained, that is, up to around 200 psi bmep.

Basic Development

Nearly all combustion system work was done on single-cylinder engines (Fig. 9). We are convinced that, provided certain precautions are taken and that care is exercised, single-cylinder development work substantially reduces development time and cost. The exact combustion-chamber configuration developed on the single was transferred to our multicylinder production engines without even minor modifications. Careful checks have shown that the optimum arrangement for both engines is the same.

---

[8] "High-Speed Compression-Ignition Engine," by C. B. Dicksee. Pub. by Blackie & Son, Ltd., London, 1946.
"Influence of Valve Port Design on Volumetric Efficiency of Compression-Ignition Engine," by C. B. Dicksee. Proceedings of Institution of Mechanical Engineers, Vol. 160, 1949.

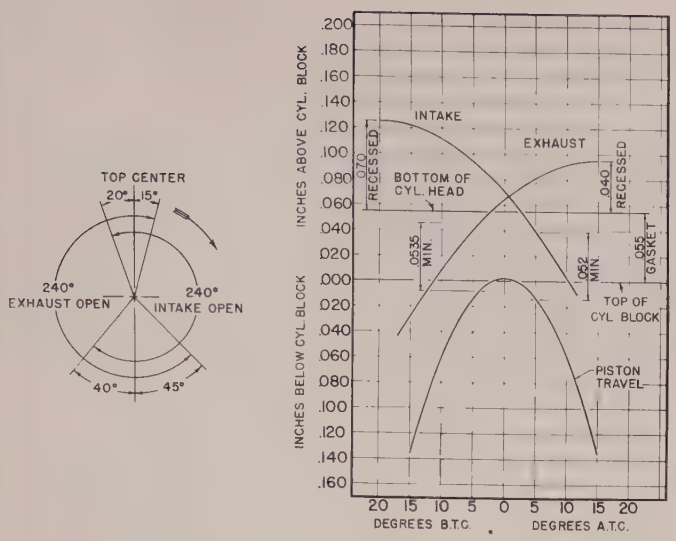


Fig. 8 — Valve timing and valve-to-piston clearance

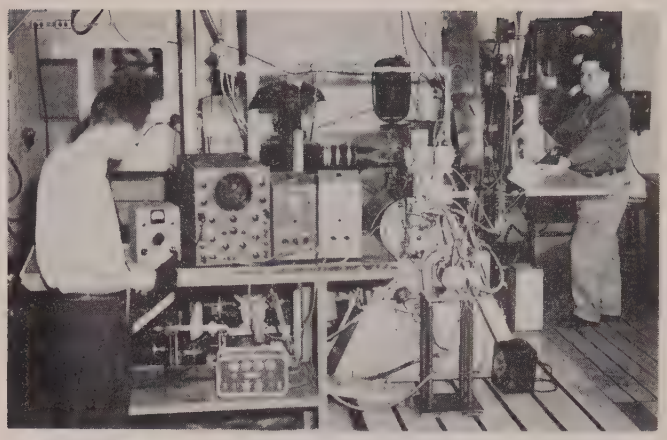


Fig. 9 — Single-cylinder engine test installation

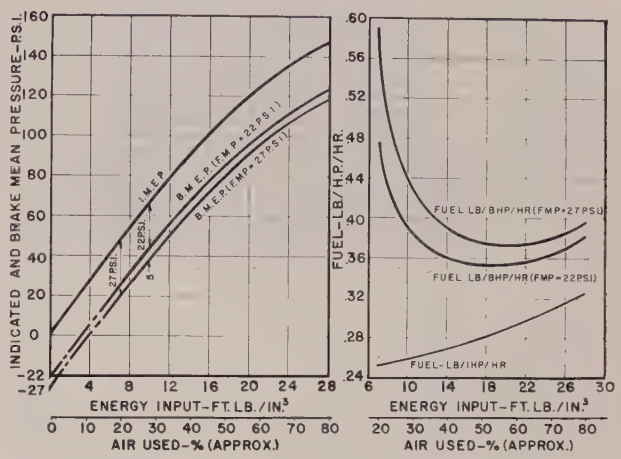


Fig. 10 — Method of evaluating indicated performance

Such agreement may be partially due to the fact that all our ports are identical, although some are mirror images of the others. We do not use siamesed intake ports.

Obviously, manifolding can have a substantial effect upon airflow pattern and can cause the multicylinder engine to perform differently from the single. If this is the case, it is advisable to correct the manifold design of the multicylinder so that its effect becomes equal to that of the induction system of the single, rather than to tailor airflow and injection to a not quite satisfactory manifold.

Only after developing all manifolds which are required for different engine applications so that they act exactly alike, can they be used interchangeably without necessitating changes in the swirl-producing arrangement.

Some engineers are inclined to doubt the practical value of single-cylinder work, probably because it is difficult to project future multicylinder performance from single-cylinder data. When it is realized, however, that necessary development changes can be made on a single in less than one-third of the time needed to perform the same work on an engine with six cylinders, it becomes evident that any time required for additional analysis is well spent.

The difficulty in corelating data is due only to the uncertainty concerning operating friction. Neither motoring tests nor the cutout method nor any of the theoretical approaches, like the Willans method, give sufficiently reliable data. The problem would be solved if indicated performance could be evaluated without the process of adding measured or calculated friction to effective performance.

A method permitting one to do just this has been developed by C. B. Dicksee and has been used by us for several years (Fig. 10). This method consists of plotting specific energy input versus brake mean effective pressure. The slope of the resulting curve provides a measure of the results obtained in return for the fuel used, in other words, it shows combustion efficiency. This method has never been described in the literature and as we feel that it would prove as helpful to other engineers as it has to us, we have suggested to Mr. Dicksee that a paper on this subject might be of considerable interest.

In the right-hand part of Fig. 10 the uppermost curve represents the brake specific fuel consumption of a single-cylinder experimental engine and the next curve that of the 6-cyl engine having the same combustion-chamber design, both operating at the same speed. It would be most difficult to tell from these curves whether or not the indicated specific fuel consumptions of both engines are the same. However, if it is found, after replotting this information as energy input per unit of displacement against brake mean effective pressure, that the resulting curves are parallel, it follows that both engines make the same use of the fuel supplied. The vertical distance between the curves then represents the difference in operating friction. The indicated mean pressure curve must obviously pass zero so that, if the bmep curve is moved into this position, the indicated pressure is obtained. It must be left to Mr. Dicksee to give further information about this method and to furnish the theory on which it is based. Our main interest is that it really works in practice.

We also found that the independently variable oil supply to the crank bearing of the C. U. E. engine and its variable temperature provide the means for doing some very useful development work on crank bearings and pistons. The piston configuration, worked out by this method, proved completely satisfactory in the multicylinder engine.

The first design, based on reasoning already de-

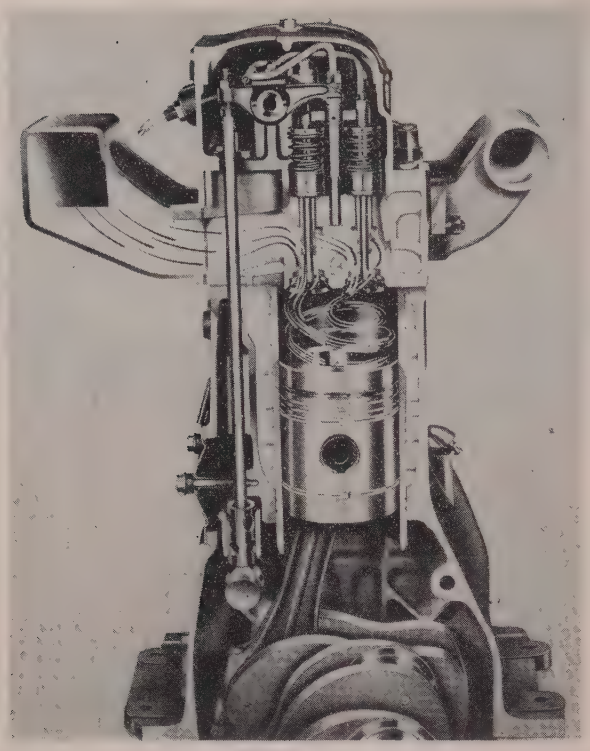


Fig. 11 — Piston configuration

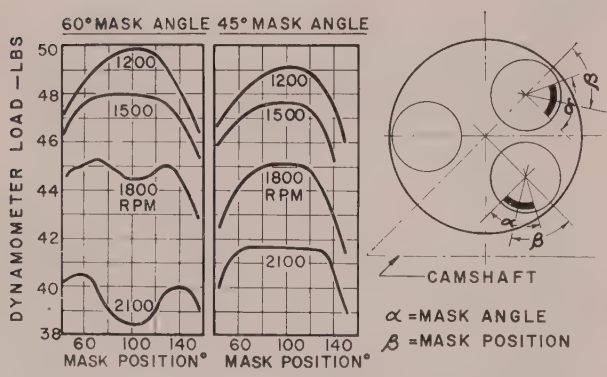


Fig. 12 — Torque versus swirl, fixed rack and fixed timing

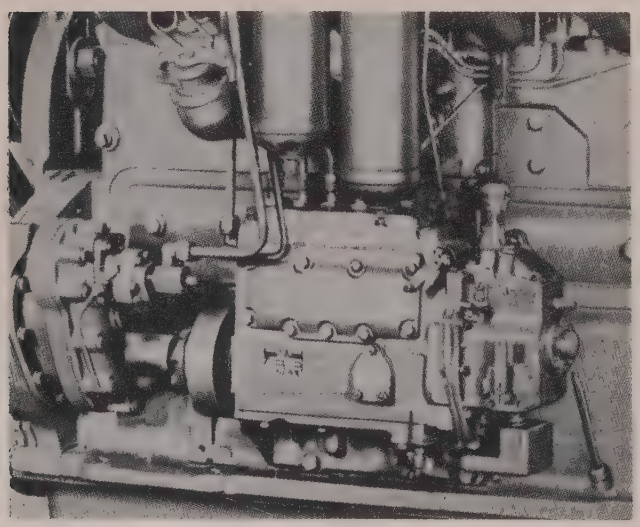


Fig. 13 — Injection pump

scribed, had piston cavity proportions as stated previously, had four nozzle holes of 0.013 diameter, 150 deg spray cone angle, 75 deg mask angle, and 15.6/1 compression ratio (Fig. 11).

When development was finished, the nozzle hole size had increased to 0.0138 for the naturally aspirated engine and to 0.0158 for the turbocharged one. This may not appear to be a great improvement over the minimum diameter of 0.012 in. which we had specified. If it is realized, however, that ease of nozzle maintenance is related to nozzle hole area, one sees that the improvement is in the order of 30 and 70%, respectively. The 150-deg spray cone angle proved to be right. The mask angle, however, was reduced to 45 deg.

Finding the proper mask angle and mask position, in other words producing the correct initial swirl rate, is the most time-consuming work. While all other variables are maintained constant, swirl alone is altered. Fig. 12 shows a pair of typical mask rotation curves for 60- and 45-deg mask angle, respectively. By "mask angle" is meant the angle contained between the valve radii which pass through the outer corners of the mask. "Mask position" is the angle between the centerline of the mask and a bore radius laid through the axis of the valve. In "zero position" the mask is closest to the cylinder bore, where it is least effective.

It is well to remember that even a maskless intake valve, placed in its customary position so that the periphery of the valve is close to the cylinder wall, is already masked very effectively over a substantial part of its total area. The valve mask, thus, merely adds to the effect of the cylinder wall. The 60-deg mask shown here is useless as its optimum positions are different for every engine speed. Furthermore, the 1800-rpm and particularly the 2100-rpm torque curves show that overswirling occurs.

The 45-deg mask is a far better match as the best mask position is the same for all speeds. In addition, performance is quite constant over a wide range of positions. This means that neither unavoidable play in the locating mechanism, nor slight deviations in the shape of the intake ports, will have an adverse effect on performance.

If no mask were used, performance would be the same as if the mask were placed on "O" position. The trend of the curve indicates clearly that this would not produce the desired performance. The mask, therefore, is necessary in order to permit the use of the large port areas which are needed to obtain high volumetric efficiency.

It was mentioned before that, in a combustion system, there is no such a thing as an "independent" variable. Any slight change in one requires variations in many of the others. Thus, after the proper swirl rate has been established for an existing injection configuration, spray characteristic, and combustion chamber, proportions have to be varied to ascertain that the optimum result has been reached.

We had developed a very satisfactory arrangement for the original 15.6/1 compression ratio when we discovered that this ratio could be reduced without any apparent adverse effect on starting. As the lower compression ratio promised lower cylinder pressures, the work on injection, swirl, cavity proportions, and the rest, had to be repeated. Finally,

we arrived at a design having a 14.5/1 compression ratio.

Injection System

The injection pump (Fig. 13) is made for us by American Bosch. Its delivery valve has as small a dead volume as can be obtained. Pump lubrication is part of the engine lubricating system. Oil supply and return passages, from the cam compartment as well as from the governor housing, are incorporated in the pump mounting bracket so that external lines are avoided.

We consider the engine-lubricated pump an absolute necessity for construction machinery. If the pump were self-lubricated, it would be just about certain that dirt would be introduced into the pump when its oil level is checked or replenished.

The oil returns are so arranged that proper pump lubrication is maintained even when the pump is inclined 45 deg in any direction.

A small amount of fuel oil must be permitted to pass between plunger and barrel to lubricate these parts. The quantity of this leakage fuel is so small that it is tolerated by a self-lubricated pump. If the pump is engine lubricated, the leakage fuel must be prevented from mixing with the lubricating oil. The ducted plungers shown in Fig. 14 serve this purpose. The pressure gradient existing along and between the mating faces of a plunger and barrel creates a substantial hydraulic pressure toward the lower end of the assembly. This pressure returns the "leak-down" to the pump gallery.

Plunger diameter determines which of the two arrangements is used. The barrel wall thickness which goes with a 10-mm plunger diameter (B) is sufficient to accommodate a single axial duct but ducting is required in both barrel and plunger when the diameter of the latter is 11 mm or larger (A).

During engine development one must be able to obtain reliable information quickly on certain parameters of injection equipment performance. Some of these are:

1. Quantity delivered by a certain pump, line, and nozzle combination.
2. Relation between a fixed point in the cycle, such as pump inlet port closing, and the actual start of injection, duration of injection, and presence or absence of secondary injections.
3. Fluid pressure during injection.
4. Nozzle needle lift.
5. Rate of injection.

Fuel quantity per pump stroke is derived from the test logs but the other items of information are not obtained as easily.

Fig. 15 shows a special injector assembly which enables us to get all the data listed under 2, 3 and 4 on the operating engine.

Start, duration of injection, and secondaries, if any, are obtained by indicating the needle lift through a transducer which is mounted on an extension of the spindle rod. We have used electromagnetic as well as condenser-type pickups but found the linear differential transformer most reliable. A network consisting of a 40,000 cps carrier, rectifier, filter, and amplifier produces a single-line trace on the scope.

Fluid pressure is indicated simultaneously by a crystal pickup loaded by a plunger which is lapped into a bore at a right angle to and intercepting the nozzle passage. Thus, the pressure is measured as close to the nozzle as possible. To give us an additional check on the electronic measurement, a mechanical maximum pressure indicator is also connected to the nozzle passage. The electronic pressure pickup does not add any volume to the nozzle assembly and the additional volume due to the maximum pressure indication is less than 20 cu mm, an insignificant increase in system volume. In order to keep the volume this low, we built the nozzle shown in Fig. 16. So far as we know, this is about as small a differential nozzle as has ever been made, having a needle lift of 0.001–0.0015 in.

Rate of delivery is usually measured by means of a "Mechanical Stroboscope" of some sort. One type

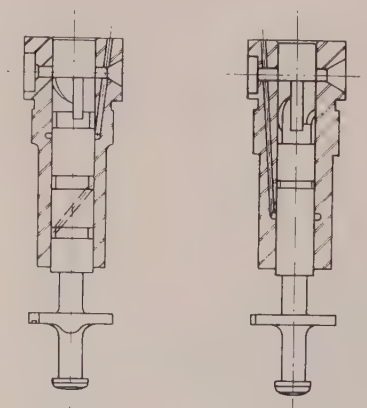


Fig. 14 — Ducted plungers

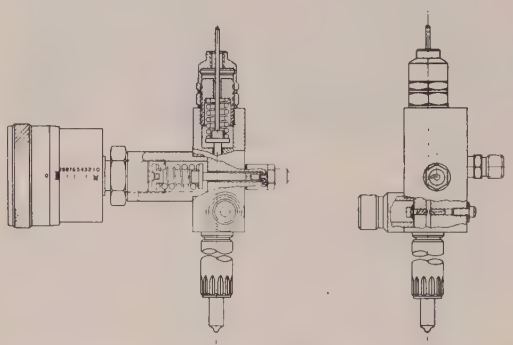


Fig. 15 — Special injector

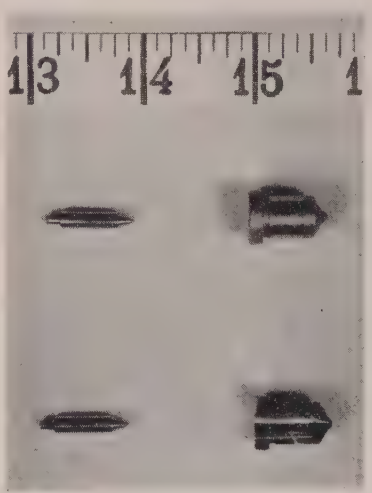


Fig. 16 — Differential nozzle

is shown in Fig. 17. It consists of a drum synchronized with the injection pump cam, which drum has a slot 1 deg wide. The machine is operated for a certain number of strokes and the fuel quantity collected by the slot is measured. By moving the slot in relation to the cam in 1-deg steps and repeating this process, the rate of fuel delivery—that is, the discharge per degree of cam rotation—can be found. This type of rate machine gives fairly reliable results but each test requires considerable time.

The chandelier machine, which has a number of compartments around the periphery of the drum, produces information in less time but its information becomes more unreliable as the pump speed increases. Furthermore, all mechanical rate machines have in common the fact that the information obtained by them cannot be recorded automatically.

The additional pickup visible in Fig. 18 takes the place of the above mentioned rate machines. It indicates the force of impingement and provides a measure for the rate of flow, as is shown by the following analysis.

The fuel stream impinging against the face of the pickup is forced to change its direction. The reaction on the pickup must be:

$$F = M \times \Delta v \tag{8}$$

where:

M = Mass flow per unit of time
Δv = Change in jet velocity

If the pickup is placed at 90 deg to the stream, the latter's momentum in the original direction is destroyed so that:

$$\Delta v = v$$

where:

v = Jet velocity at nozzle

Therefore:

$$F = M \times v$$

if:

Q = Volume flow per unit of time
ρ = Density

$$M = Q \times \rho \tag{9}$$

with:

C = Orifice discharge coefficient
A = Orifice area,

$$Q = CAv \tag{10}$$

or:

$$v = \frac{Q}{CA} \tag{11}$$

Combining (8), (9), and (11):

$$F = \frac{Q^2 \rho}{CA} \tag{12}$$

$$Q = \sqrt{\frac{FCA}{\rho}} \tag{13}$$

By combining all constants into K and expressing rate of injection in terms of crankshaft degrees, we can write:

$$R = K \frac{\sqrt{F}}{N} \tag{14}$$

where:

N = Engine speed, rpm
R = Volume flow per deg crankshaft rotation
K = Proportionality factor

If the volume of pump delivery per cycle V is known, K can be found from the relation:

$$V = \int_{\alpha_1}^{\alpha_2} R d\alpha = \frac{K}{N} \int_{\alpha_1}^{\alpha_2} \sqrt{F d\alpha}$$

or:

$$K = \int_{\alpha_1}^{\alpha_2} \frac{VN}{\sqrt{F d\alpha}} \tag{15}$$

where:

α_1 = Crank angle at beginning of injection
α_2 = Crank angle at end of injection

By dividing the impingement force trace into 1-deg intervals and summing the square root values of F, the value of K is found for the particular nozzle. K is valid for all subsequent tests of this nozzle and, when introduced into (14), gives the volume flow per degree crank angle.

In practice, the photograph of the crystal pickup trace is measured at 1-deg intervals and rate of discharge is obtained from a simple computor manipulation. Fig. 19 shows a comparison between rates of discharge obtained by mechanical and pressure pickup methods. We feel that agreement is quite satisfactory, particularly as it is by no means certain that the information provided by the mechanical method represents the whole truth.

Time saving is not the only advantage: It is clear that this device permits simultaneous measurement of the individual rates of discharge from each hole of a multihole nozzle even when injecting into a pressurized atmosphere. This is information which it is practically impossible to obtain from a mechanical rate machine.

Data Presentation

Some performance data and other test results are given later in this paper. Where appropriate, we intend to state what type of instrumentation was used in obtaining such data, as omission of this information makes it rather difficult for the reader to compare published test results with those which he has obtained himself.

All our test cells (Fig. 20) are equipped with filtering-type smokemeters and the recording of smoke data is part of our standard test procedure. The filter tape trace is evaluated by a light transmission photoelectric meter. However, as a standard yardstick for measuring smoke intensity does not as yet exist and as we have not been able to produce photographic slides which would clearly show the mean-

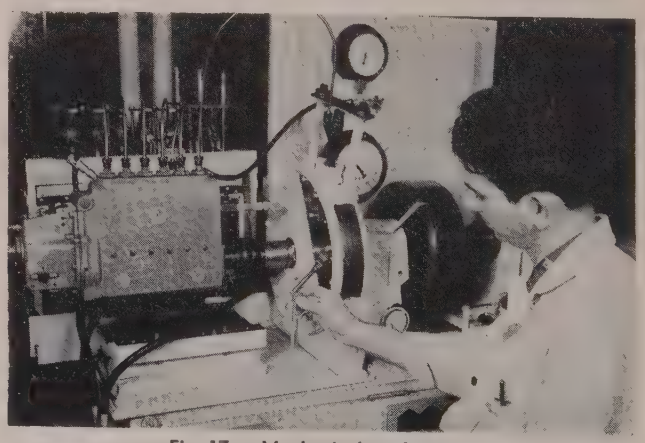


Fig. 17 — Mechanical stroboscope

ing of our smoke numbers, the following description attempts to clarify our meaning:

No. 20 smoke is completely invisible.

No. 25 corresponds to the point where the smoke becomes just discernible against a white background.

No. 35 is a visible light smoke about equal to 50% light reflectance by the CRC method.

No. 50 would still safely pass the scrutiny of a Los Angeles County Smoke Inspector, whereas

No. 60 would probably not do so if it were maintained over any appreciable distance.

No. 70 is fairly heavy smoke, probably about 25% CRC, but is still well below the extra heavy smoke put out by an engine which is being pushed to its limit. A smoke number of 80 would approximate this condition but we cannot state this definitely as readings above 75 become rather unreliable.

All performance data, unless specifically stated otherwise, are observed, no correction factor being used. The air-inlet temperature is controlled to 85–90 F and the barometric pressure is that prevailing in the test cells, namely 0.1–0.3 in. water column below ambient, maintained at this value by exhaust fans. Usually we prefer to use observed data in our development work rather than to attempt corrections.

We have a method for obtaining reliable corrected data, based on air utilization, which gives excellent results but which is rather cumbersome and time consuming.

It will be noted that some performance figures, shown in our sales literature, are inferior to those given in curves contained in this paper. This is due to the fact that the latter were obtained from fully run-in laboratory engines whereas the sales literature curves show the minimum performance of production engines.

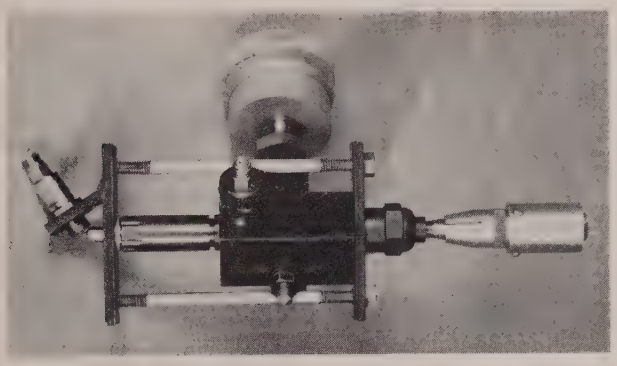


Fig. 18 — Pickup which measures rates of flow

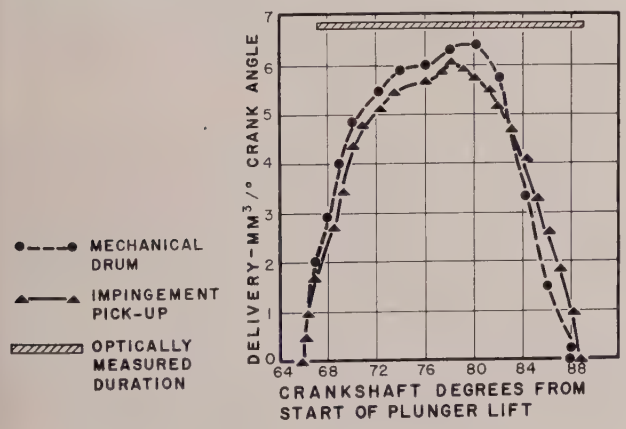


Fig. 19 — Rate and duration of injection by three methods

Some Design Details

It was mentioned earlier that cylinder distance was retained in the new engines, but no major elements of the new engines are interchangeable with those of the older ones.

Crankshafts

The crankshaft received our particular attention. Fig. 21 shows the results of bending fatigue tests carried out with the old 844 crankshaft as well as with two new crankshafts. These tests were performed in a resonant fatigue testing machine which is calibrated statically for each type of test specimen. The specimen is regarded as broken as soon as a change occurs in magnitude and shape of the signal produced by two of the strain gages, that is, at the very onset of a fatigue crack.

As mentioned later, old style crankshafts had to be used in the first batch of experimental engines. Although they did not fail, some of these crankshafts showed that they were operating near the

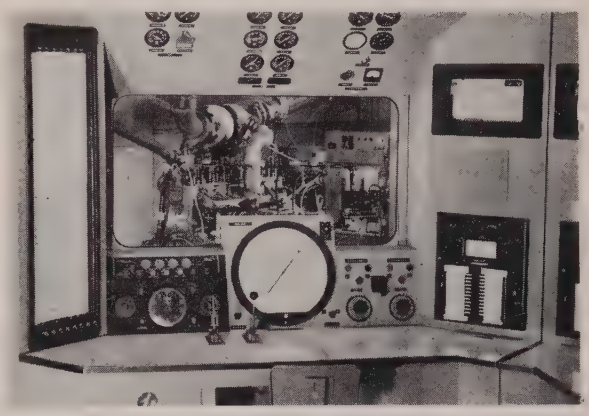


Fig. 20 — Test cell

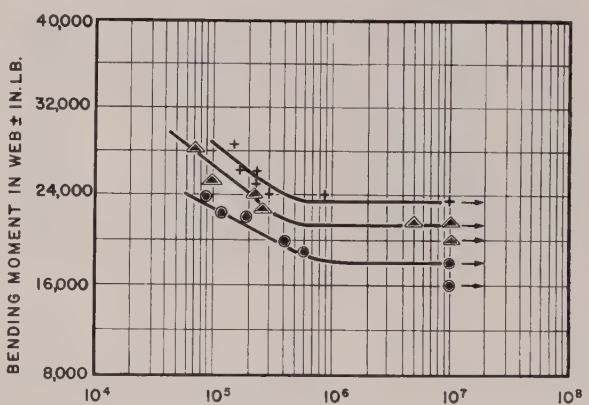


Fig. 21 — Bending fatigue tests

Fig. 22 — Crankshaft

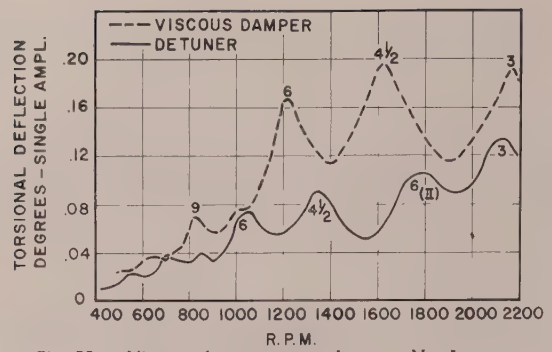


Fig. 23 — Viscous damper versus detuner, No. 1 power curve ("21000" series)

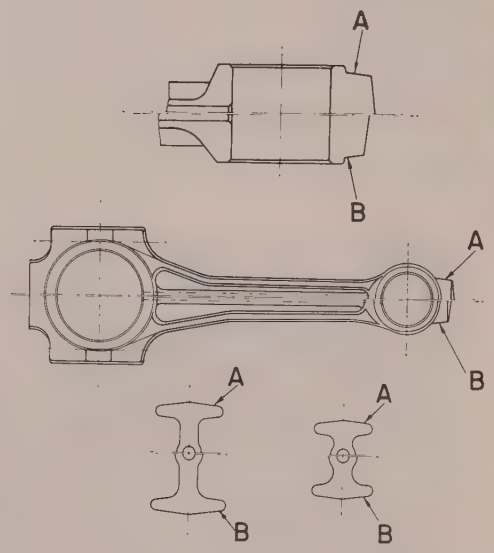


Fig. 24 — Connecting rods

fatigue limit. Thus, the 28% higher fatigue strength of the new shaft (Fig. 22) provides a comfortable margin of safety.

The condition was further improved by a substantial reduction in torsional amplitudes, resulting from the use of a detuner in place of the former damper. Fig. 23, obtained from a C.E.C. torsiograph and meter at full No. 1 curve output of the turbocharged engine, shows the comparison.

It should be pointed out that the old arrangement with its 0.20-deg single amplitude was entirely satisfactory. Nevertheless, the amplitude reduction of some 50% is certainly worthwhile.

An incidental benefit, as yet not proved by endurance tests but which can reasonably be expected, is increased belt life. In the old arrangement, the belts were driven by pulleys keyed to the crankshaft, so that they participated in crankshaft vibrations. Now, however, the belt grooves are incorporated in the inertia mass, which runs at substantially constant angular velocity.

SAE 1046 is used for this forging. It is heat-treated to 248/277 Brinell and all bearing surfaces are induction hardened to Rockwell "C" 48/52. Rolling the fillets and changing the location of the oil supply hole in the crankpin has increased the fatigue strength of the shaft beyond the values shown in the graph.

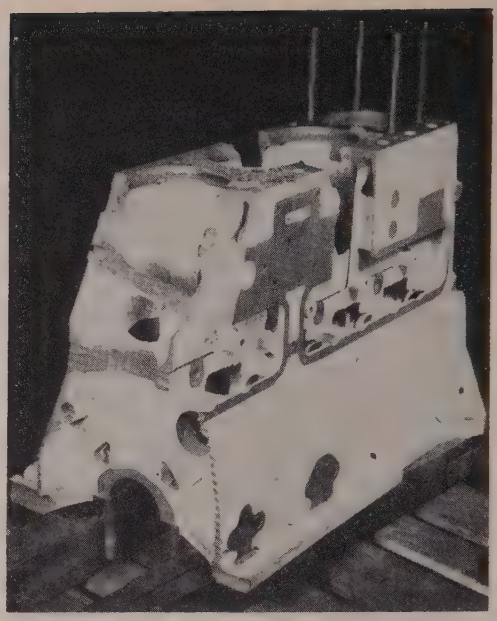


Fig. 25—Cylinder block

Connecting Rod

Although they stood up better than expected, connecting rods borrowed from the older engine model for experiments with the new combustion system confirmed that a stronger rod was needed. This had been anticipated by design (B), which is shown in Fig. 24 in comparison with the rod formerly used (A). It will be noted that the shaft sections have been increased substantially and that filleting, particularly of the upper end, has been much improved. The rods are forged from SAE 1041 steel and are heat-treated to 207/241 Brinell.

Cylinder Block

In designing the cylinder block, particular attention was paid to proper support for the cylinder sleeves, not only by substantial curtain walls and adequate shoulders but also by stud bosses faired into these shoulders. Stud loads are transmitted to the main bearings saddles by thickened sections of the block casting which act as tension bands. The flat portions of the side walls are practically stress-free (Fig. 25).

Thread engagement of the cylinder head studs is twice the thread diameter and the stud holes are counterbored approximately to the level of the liner shoulders to avoid stress concentrations between threads and cylinder counterbores. A 40,000 psi minimum allow iron is used for the block casting.

Cylinder Head

The cylinder head (Fig. 26), in accordance with one of our original requirements, is a straightforward design. The flat underside and ample studding permitted a reduction in height of 1¾ in. compared with the cylinder head of the energy-cell engine, while still retaining more than adequate core thicknesses.

The cooling water flows across the bottom deck from the push rod toward the exhaust side. By locating a major water inlet hole directly under the intake port and the water outlet above the exhaust port, effective cooling of all critical areas is obtained

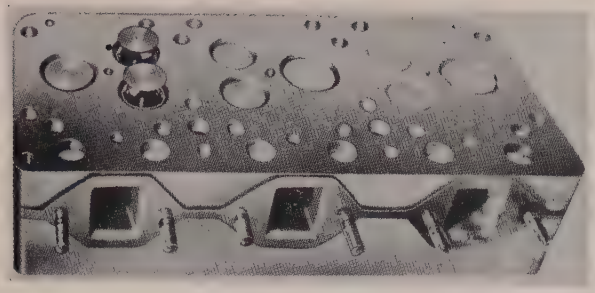


Fig. 26 — Cylinder head

Fig. 27 — Intake valve bridge

Fig. 28 — Valve and injector arrangement

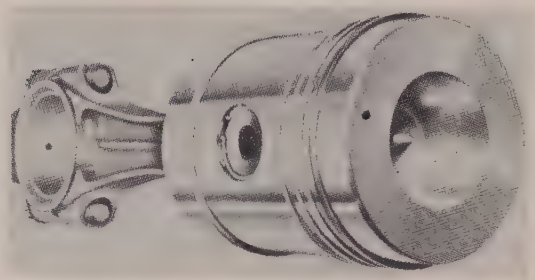


Fig. 29 — Piston and connecting rod assembly

Fig. 30 — Pistons removed from turbocharged endurance test engine

without the use of directional tubes or difficult to core passages. Stud bosses are separated from the valve ports by cores of such size that burning-in is avoided.

We have found that high stresses can be imposed on the bottom deck if the usual injector holding nuts are carelessly overtorqued. This possibility is avoided by the use of a single stud and crab, one end of which rests on the rocker arm support bracket, the other on the injector assembly. This arrangement, while providing ample nozzle gasket sealing pressure, is foolproof.

Intake valve bridges (Fig. 27) are SAE 4140 forgings, heat-treated to Rockwell C 28–32. Rocker arms are forged from SAE 1117, carburized, and hardened to R_c 56–63 on the nose end. The microstructure of the valve guide bore is limited to 5% free ferrite and must check Rockwell B 88 minimum. Rocker arm shafts are made from SAE 1045 tubing and are induction hardened to Rockwell C 55–62.

As mentioned earlier, the intake valves are nonrotating (Fig. 28). Their position is maintained by a sliding key milled into the upper part of the valve stem. A valve locator, piloted on the valve guide and fixed in a definite angular position by a locating pin in the cylinder head, permits free axial movement of the valve but maintains the "mask position." The two concentric and counterwound intake valve springs are so designed that their respective windup torques neutralize each other. Thus, the locating key is practically free from any load, particularly as the guided valve bridge exerts only axial forces on the valve stem. Aerodynamic forces on the valve mask are negligible.

The valve mask is forged integrally with the head and it is machined all over in the form of a complete 360-deg mask. The excess beyond the required mask angle is milled off so that the mask centerline coincides exactly with the key.

Piston

The aluminum piston (Fig. 29) has a Ni-resist insert which carries the keystone-type top ring. The contour of the piston cavity is such that it provides adequate heat paths from the upper portion of the cavity but slightly restricts the heat flow from its bottom. No specific arrangements for oil

cooling are made as the cooling provided by the existing splash is ample for the ratings used to date. This is borne out by Fig. 30, which shows a set of pistons removed from a turbocharged engine which had operated for 1000 hr at continuous full-rated load.

Bearings

The bearings are steel-backed aluminum. The wall thickness is 3/6 in., the aluminum layer approximately 0.020 in. Upper and lower halves of main and connecting-rod bearings are interchangeable. The set of connecting-rod bearings shown in Fig. 31 came from the same test as the pistons mentioned above.

Front Support

The engine front support (Fig. 32) appears to be so simple an item that it should not pose any problems. Nevertheless, the materials which are standard for this application were not satisfactory. A molded nylon liner, contained in a cast-steel housing, has apparently solved this difficulty completely.

Lubricating System

A schematic drawing of the lubricating system is shown in Fig. 33. As the engines must operate at angles up to 45 deg in any direction, a scavenging oil pump is used to transfer the oil from the front end of the oil pan to the deep sump when the unit is in a nose-down position. A pressure relief valve, set at approximately three times operating oil pressure, is installed in the discharge passage from the pressure oil pump to protect the system during low-temperature starts. A common header casting for oil filters, lubricating oil cooler, converter cooler, and filter bypass valve has eliminated all external oil piping in the primary system, substantially reducing oil warmup time. Oil filter cans are suspended from the header casting to avoid spillage when the filter elements are serviced. The oil cooler was placed in a vertical position to guarantee complete water drainage. A separate turbocharger oil filter is also connected to the header. (See Fig. 1.)

Rear Power Take-Off

In some applications heavy auxiliary power requirements are encountered. Space limitations sometimes prohibit driving the auxiliary equipment from the front end of the crankshaft. The arrangement shown in Fig. 35 solves this problem. The engine assembly is standard to the crankshaft flange but the rear oil seal is moved from its normal location on the rear face of the crankcase into a vertical wall of a special flywheel housing. This provides space for an extremely husky gear clamped

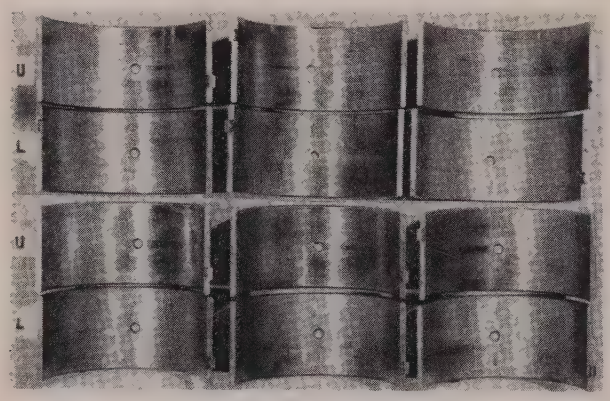


Fig. 31 — Connecting-rod bearings

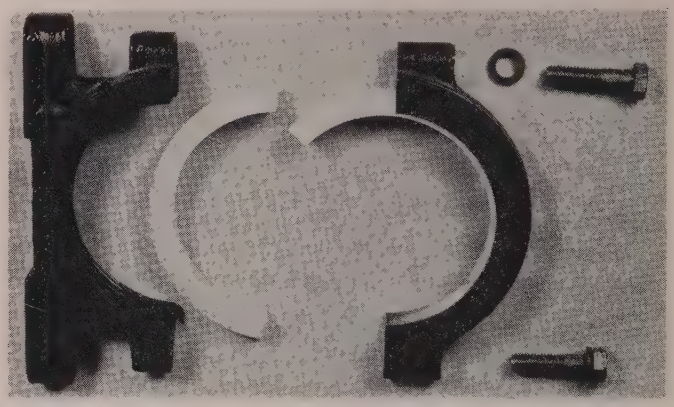


Fig. 32 — Engine front support

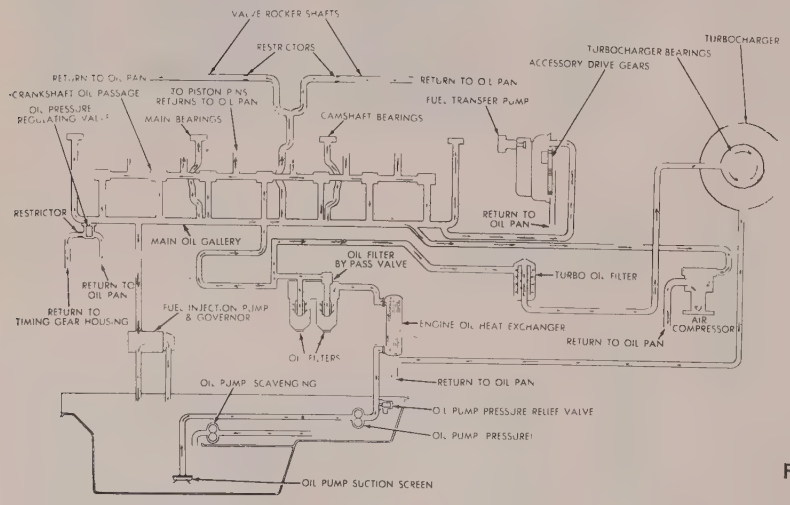


Fig. 33 — Schematic of lubricating system

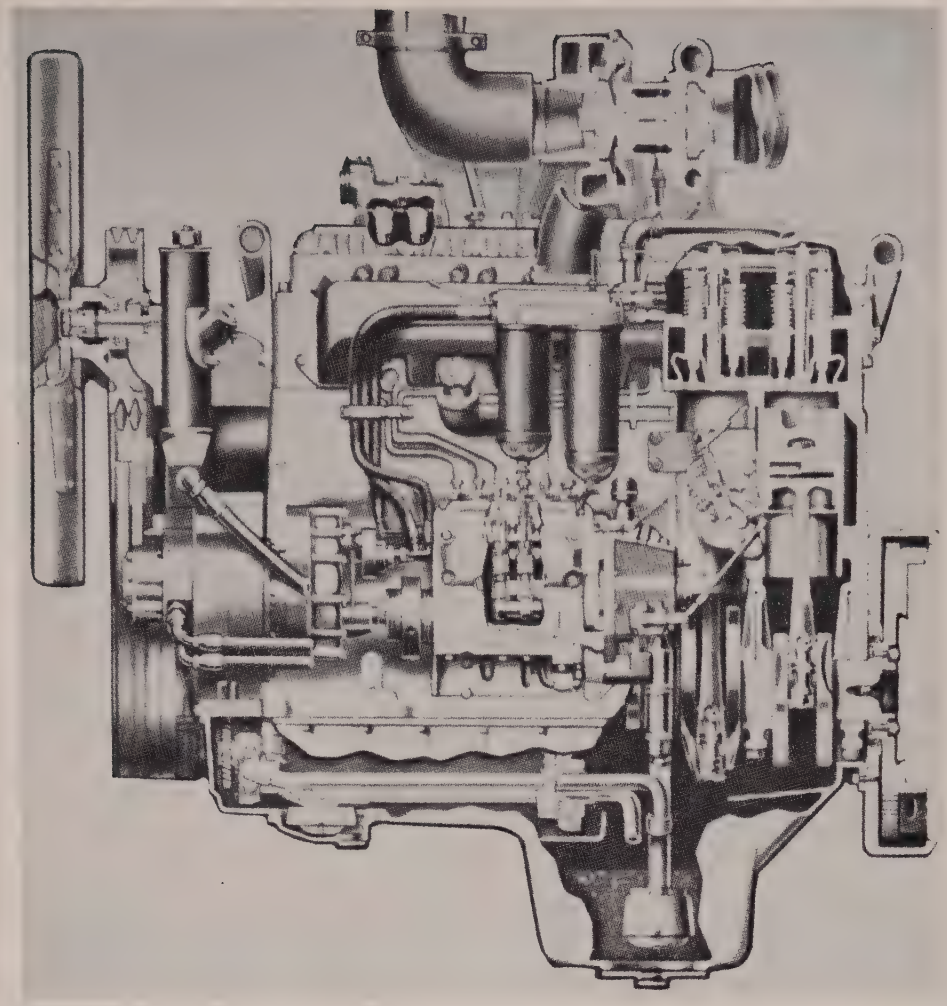


Fig. 34 — Cutaway of "21000" engine

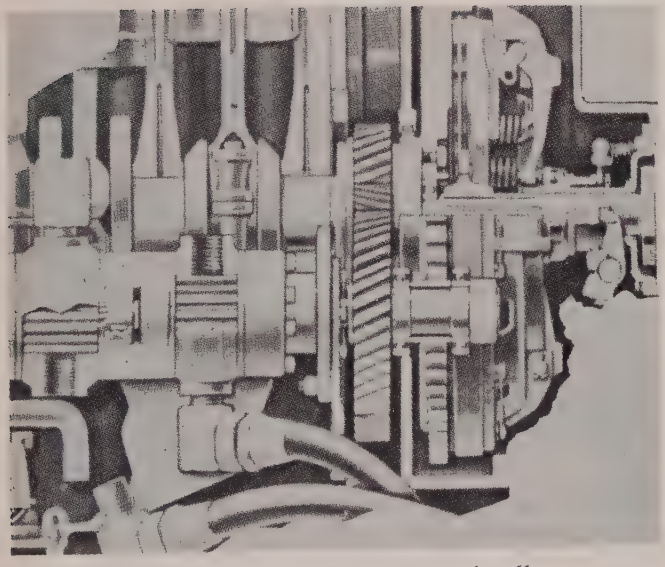


Fig. 35 — Cutaway of rear power take-off

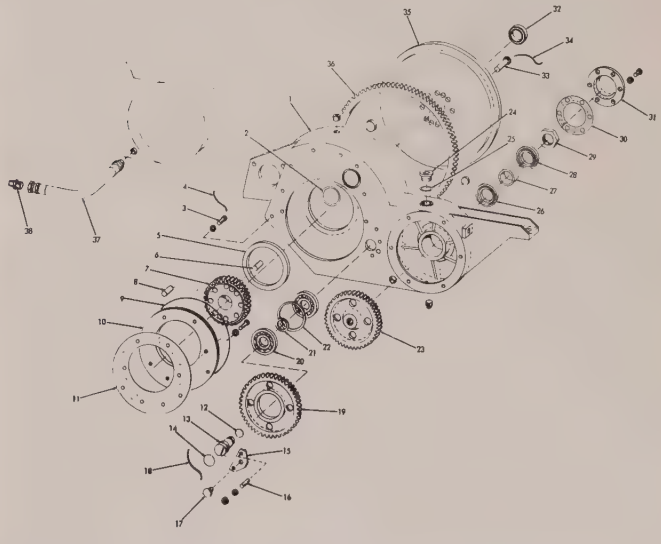


Fig. 36 — Rear power take-off, exploded view

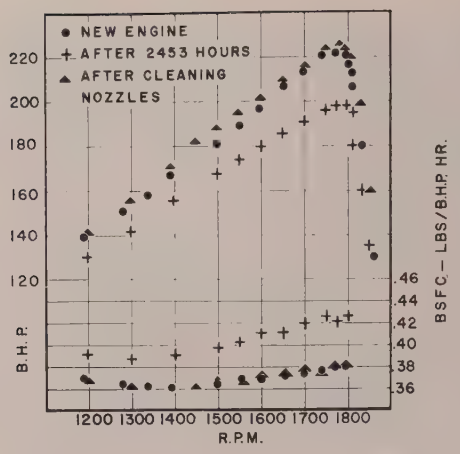


Fig. 37 — Performance curve of "16000" endurance engine

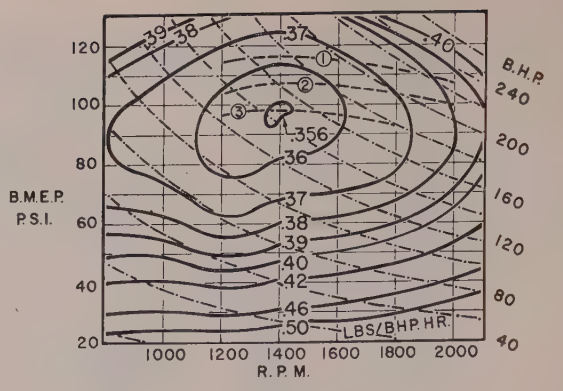


Fig. 39 — "16000" performance map

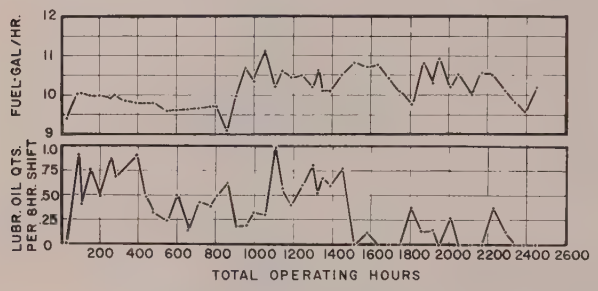


Fig. 38 — Fuel and lubricating oil consumption, Proving Ground ("16000" engine, 220 bhp at 1815 rpm)

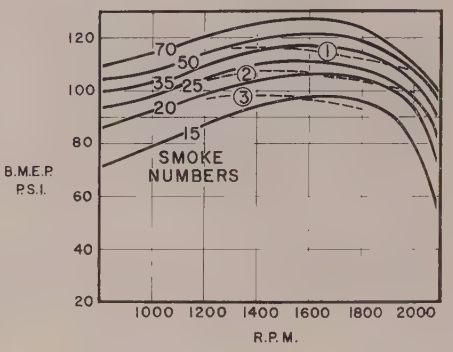


Fig. 40 — "16000" smoke number map

between crankshaft flange and flywheel. A box-shaped extension of the flywheel housing contains an idler and auxiliary drive gear, together with their shafts and bearings. This gear housing also acts as rear engine support (Fig. 36).

Testing Experimental Engines

As soon as combustion system development had been completed on single-cylinder engines, two 6-cyl units were built for performance testing. Newly designed parts were not as yet on hand so that we had to "cobble-up" these engines out of crankshafts, connecting rods, cylinder blocks, gear trains, and other parts from older engines.

It has been mentioned already that the combustion system proved to be directly applicable to the multicylinder engines so that we found ourselves ready to build endurance and proving ground test engines much earlier than had been anticipated. Rather than wait for the new parts to become available, we decided to assemble 20 engines by the same method which had been used in the first two. We were aware that some of the parts, borrowed from energy-cell-type engines, would fail. This was of no importance as long as the combustion system proved itself.

Quite naturally we had some troubles, all of them relatively minor, but quite annoying: Valve adjusting screws came loose and intake valve bridges came out of engagement. We had located the bridge only on one valve stem as it seemed reasonable that this would suffice. Apparently, however, the closing rates of two identical valves may differ sufficiently to lift the single locating cup out of engagement. When this occurs, valve stems, push rods, and other things suffer. Locating the bridge on both valve stems proved to be the remedy.

The piston-ring setup, which had looked perfect after constant-speed and cycling dynamometer tests, proved to be quite unsatisfactory for scraper service.

Overspeeding tests on engines installed in rubber-tired vehicles showed up shortcomings in valve-to-piston clearance and in valve spring design.

I would like to mention here that no design precautions can protect an engine if the driver is particularly adept with clutch and stick. When his rubber-tired vehicle is coasting downhill at full engine speed and he succeeds, by a sudden downshift, in suddenly increasing the rear axle-to-engine ratio by 2 or 3/1, then the 50 tons traveling at over 20 mph will certainly provide enough energy to overspeed the engine so that something must happen.

The organization of Allis-Chalmers is such that the testing of new engine models is done in at least two, but usually in three or four, different locations. When the Engine Division has completed the development of a new model and is satisfied with its own endurance tests, several pilot engines are turned over to our customers — namely, to those divisions of the company who build tractors, scrapers, farm machinery, and the like. These engines are not only installed in various vehicles operating on the Proving Grounds, but they are also subjected to all kinds of additional dynamometer tests as all plants, which use these engines, are equipped with engine test facilities.

As is always the case amongst relatives, these sister plants are our most severe critics and they make wholehearted efforts to find out what is wrong with what we have done. It is the duty of their test

Table 4 — Fuel Specifications

Fuel Type	A	B	C
Gravity, API	40.8	38.8	34.2
Heating Value, Btu/lb net	18,575	18,570	18,425
Sulfur, %	0.11	0.15	0.7/0.83
Distillation F			
Ibp	336	363	365
10%	388	428	434
50%	480	486	516
90%	576	567	599
EP	625	640	657
Cetane No.	51	53	42
Viscosity SSU at 100 F	35	34.1	35.5
Carbon Residue, (10% bottoms, Ramsb.)	0.07	0.1	0 25

personnel to subject all equipment to every kind of use and abuse which it might encounter in service. These engineers have learned from long experience how to torture a machine in a scientific way and whatever they cannot break, bend, damage, or destroy will withstand any treatment which it might receive in the field. It would be difficult to overemphasize what these men contribute to the quality of the finished product.

Normal development and endurance testing is done in Harvey with fuel designated A in Table 4. The other plants use Fuel B for most of their dynamometer work but Proving Ground equipment is operated on Fuel C. We believe that these fuels cover a sufficiently wide range to fairly represent the variety of fuels used in the field.

The engines of the first group were set to perform in accordance with the original power specifications. These engines operated for 600–1500 hr, that is, until the newly designed parts had become available.

Encouraged by the results obtained with the first group, it was decided to load the prototypes well beyond their respective rated output. To do this, naturally aspirated 16000 engines, intended for the HD-16 tractor, were set at 220 bhp and were installed in HD-21's. 21000 engines were opened up to deliver 300 hp instead of their normal tractor rating of 240 hp.

A fairly typical report, received from our tractor plant, summed up one of the tests as follows: "Engine 16000, Serial No. . . . . , was removed from the tractor after 2453 hr of operation. Regular filter and oil changes had been performed. The engine was subjected to a dynamometer test but nozzles had to be cleaned before 'as new' performance was obtained. The enclosed curve sheet shows performance as received from the proving ground and after cleaning nozzles." (See Fig. 37.)

It would seem that having to clean nozzles after nearly 2500 hr, should not be cause for complaint, particularly when the engine had been operating at 122% of its normal tractor rating and had been using 0.83% sulfur fuel. It is interesting that this sort of thing can be done but it should certainly not be regarded as the proper way of treating an engine, since normal procedure calls for inspecting nozzles about every 1000 hr.

In proving ground work the loading is such that the fuel rack is tight against the smoke stop for more than 80% of the operating time. Fig. 38 bears this out. It can be seen that, during the last 1500 hr of operation, the hourly fuel consumption averaged 10½ gal/hr, equivalent to about 196 hp-hr. Thus, the engine was operating at an average load factor of 109% of normal tractor rating. The change in the character of the hourly consumption curve at 800 hr is due to the fact that up to this time the tractor was towing a stone boat. After 800 hr it was put to heavy dozing.

Lubricating oil consumption is shown also. An explanation is in order as we do not want to create the impression that this engine does not use any lubricating oil at all. During the first 1500 hr, the lubricating oil content was brought up to the "full" mark after every shift and the total of these additions during each 100 hr period divided by 8, is plotted as "qt per 8 hr shift." At 1500 hr it was decided to add oil only when the oil level was getting near the "low" mark. "O" consumption, then, means that no topping-off was needed during that particular 100-hr oil change period, so that the consumption was less than the 5 qt differential between the "high" and "low" marks on the dipstick.

Before production was started, 22 engines had been used in prototype testing, for periods of 2400–7000 hr and totaling some 100,000 engine hr.

No major change was found to be required but several small items had to be improved as I have indicated before. All in all, we are quite well satisfied with the results obtained.

Performance

Whether or not the original design requirements have been met can best be judged from performance data.

Fig. 39 shows fuel economy of the 16000 engine. Curves of constant specific fuel consumption are plotted against bmep and rpm. Also shown are the rating curves for commercial engines. These curves are defined as follows:

Curve 1 — *Full Power Setting:* Represents the power available at full-throttle for engines to be installed in off-highway rubber-tired equipment, for stand-by service or similar applications. Production engines will produce horsepower within 5% of values shown.

Curve 2 — *Intermittent Duty:* Recommended power to be used for applications having varying loads and speeds with full power being required for short periods.

Curve 3 — *Continuous Duty:* Recommended power to be used for driving sustained full loads for 24 hr per day operation.

Performances shown in this as well as in later illustrations were obtained under ambient conditions as already stated, and with standard air filters and laboratory exhaust system connected. The exhaust back pressure varies from 1-in. water column at low engine speeds and loads to 11 in. at the high end of the performance curve.

Constant smoke densities are plotted in Fig. 40 against the same coordinates as were used in the previous illustration. Smoke No. 25, which defines the limit of completely invisible exhaust, lies higher than Curve 2 over the latter's full operating range. By referring to smoke number definitions given previously, it can be seen that exhaust color is quite acceptable in the case of Curve 1 also, as it approaches Smoke 50 only at the very top and bottom of the speed range. As a matter of fact, the loca-

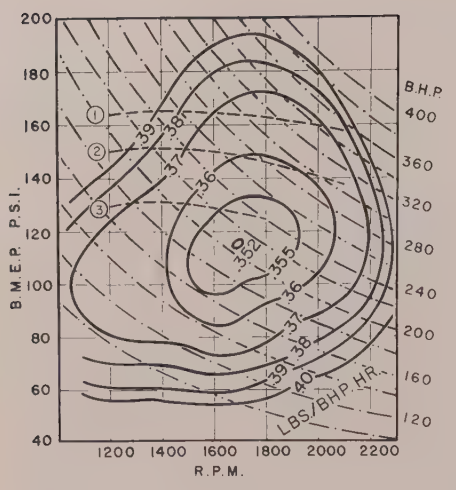


Fig. 41 — "21000" performance map

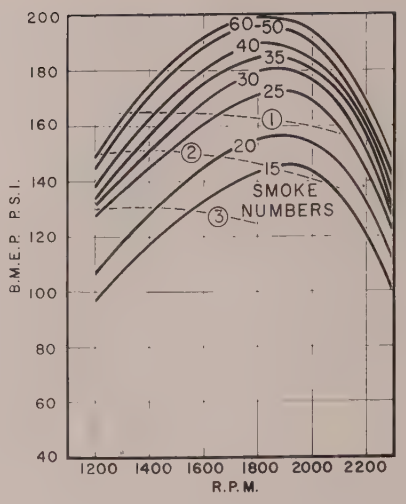


Fig. 42 — "21000" smoke number map

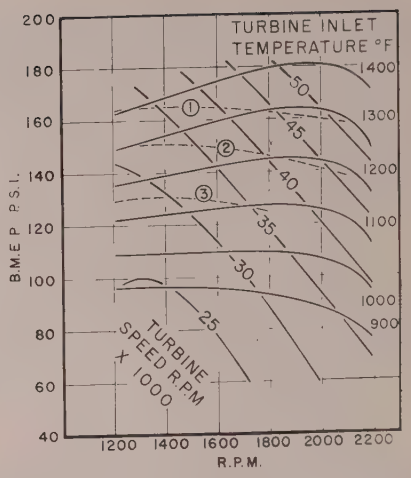


Fig. 43 — "21000" turbine speed and exhaust temperature map

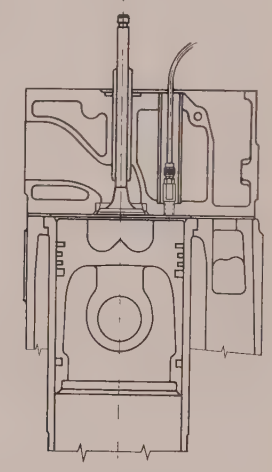


Fig. 44 — Installation of pressure transducer

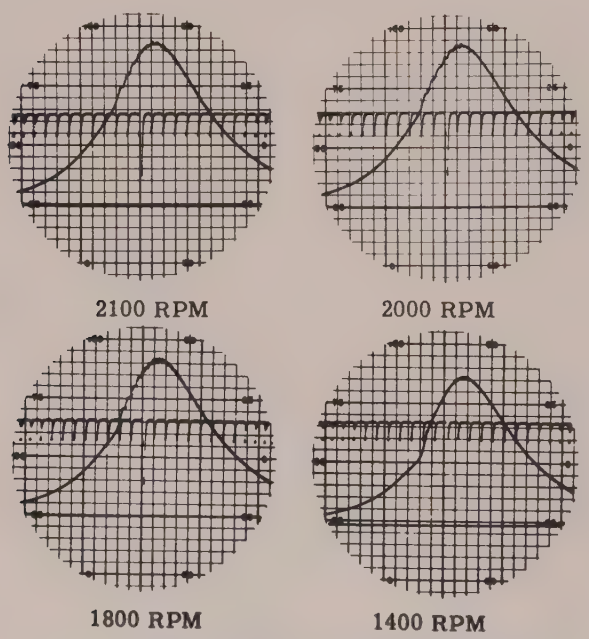


Fig. 45 — Traces of "21000" engine operating at 165 psi

tion of Curve 1 is determined by the requirement that its full operating speed range must lie below Smoke 50. As the smoke contours have a higher "hump" than the torque curves, their curvature automatically places most of the operating range below Smoke 35, that is, within the region of very light haze.

Fig. 41 gives fuel consumption contours for the 21000 engine as equipped with the "standard" turbocharger. This unit is matched to cover most of the usual applications having governed speeds from about 1700–2000 rpm.

A different turbine configuration is used for "slow speed" applications in the 1200–1500 rpm governed range. The map of the last mentioned model is practically the same as the one shown, except that the contours are moved toward the left by 400 rpm and that they are lowered by about 18 psi bmep. This places minimum specific fuel consumption at 1300 rpm and 100 psi bmep. The peak of the 0.37 contour occurs at 1400 rpm and 160 psi bmep.

Again, another turbocharger arrangement translates the contours upward and toward the right when higher maximum output is needed. So far we have merely established this fact experimentally but have not pursued the development as we concentrated our efforts on engines particularly suited to the requirements of construction machinery. Until recently, perfecting the tractor and scraper engines was more important to us than achieving maximum output figures.

Smoke data for the same "standard" engine are given in Fig. 42. Clearly, smoke does not limit the full power rating in this case as even Curve 1 is well below the visibility limit for more than half of the speed range. Satisfactory acceleration and reasonable exhaust color during acceleration determine what ratings can be chosen.

Turbine speeds and turbine inlet temperatures are shown in Fig. 43. They are such that the "standard" engine can, without any adjustment, operate up to elevations of 10,000 ft on the setting corresponding to Curve 3.

Heat Rejection

The specific fuel consumption figures just shown make it quite unnecessary to discuss in detail heat

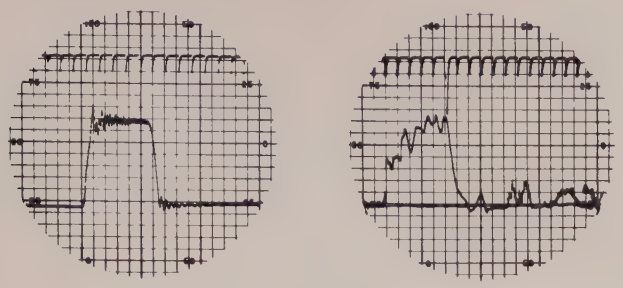


Fig. 46 — Needle lift and fuel pressure at 1800 rpm and 165 psi bmep

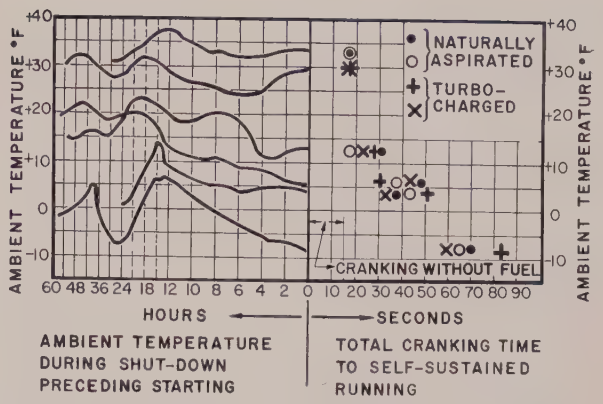


Fig. 47 — Low temperature starting without starting aids

rejection to the cooling water. One datum point for each of two engine models will suffice, as shown in Table 5.

The increase in specific fuel consumption of either engine, between Curve 3 and Curve 1, is less than 5%. Heat rejection rates of less that 27 btu/bhp min at maximum output and speed are, therefore, quite reasonable. Thus, the cooling requirements of the new engines are more than one third lower than were those of our earlier engines.

Table 5 — Heat Rejection of New Engines

Model	Bhp	Rpm	Bsfc lb/bhp hr	Heat Rejected to cooling Water, btu/bhp min
16000	180	1800	0.368	25.6
21000	240	1800	0.353	25.2

Cylinder Pressures

Before speaking about cylinder pressures, I would like to plead that a more realistic attitude towards this subject be taken by the engine industry. Much too often one encounters published information on cylinder pressures and related events which cannot be reconciled with the laws of thermodynamics. These laws, it is well to remember, show clearly that an increase in mean effective pressure must be accompanied by increased peak pressure if specific fuel consumption is not to be sacrificed. If the latter were done, rising exhaust temperatures would quickly put a stop to further improvements in mean effective pressure. Thus, there exists a definite relationship between mean effective pressure, peak pressure, and fuel consumption, a relationship that cannot be altered by glowing phrases sometimes contained in sales literature.

Fig. 44 shows the installation of the SLM pressure transducer which we use with a Kistler Piezo-Calibrator and a Dumont 322-A oscilloscope to obtain pressure diagrams. Needle lift and line pressure traces are produced with the instrumentation previously described and an Electroproducts Synchromarker supplies the top dead center mark and 5-deg pips.

The calibration in Fig. 45 is such that each square on the screen represents 100 psi on the Y axis and 5-deg crank angle on the X axis. The traces are from a standard 21000 engine operating at a constant mean effective pressure of 165 psi at the speeds indicated in the illustration.

The traces show that cylinder pressure increases slightly with engine speed. This increase is largely offset by the rise in reciprocating forces so far as bearing loads are concerned. As a matter of fact, the magnitude of gas loads is of no particular importance provided, of course, that the engine is properly designed to carry the resulting forces.

Shock loads are a different matter as they are the major cause of fatigue failures. Since a rapid change in rate of pressure rise will produce a severe shock, the shape of the transition between conpression and firing pressure is of greatest importance. Ricardo has shown that a sudden change in the direction of the pressure trace invarably produces a rough engine; but that, if this transition follows a smooth curve, the engine will be free of diesel knock. In other words, it will not be subject to shock loads.

At speeds from 1800 rpm upwards, the transition is so gradual that it is very difficult to distinguish the firing point. At lower engine speeds the point of ignition can be seen, but the change from compression to firing remains remarkably smooth.

Although injection timing is fixed, peak pressure occurs, at all speeds, at practically the identical crank angle of 7 deg atdc. Clearly then, initial swirl and secondary air movements, together with a properly matched injection system, are doing their job.

Fig. 46 shows needle lift and fuel pressure at 1800 rpm and 165 psi bmep. Nozzle opening at 20 deg btdc is indicated by the trace. The opening point can also be seen on the pressure trace and it shows that the dynamic opening pressure is some 1500 psi higher than the statically adjusted opening pressure of 2500 psi. The tdc mark of the pressure trace is moved 15 deg to the left of the vertical centerline of the scope screen so that the residual pressure waves, following closing of the needle, can be observed. The highest value of wave pressure barely reaches half the nozzle opening pressure, showing that no secondary injections occur even at the condition shown, which corresponds to the highest fuel quantity delivered in any application.

Cold-Weather Starting

It will be remembered that "better than average cold-weather starting" was one of the original re-

quirements. Cold-room tests indicated that the engines possessed very satisfactory starting ability, so much so, that we made it a rule to always crank a cold engine for 15 sec, without admitting fuel, to make sure that lubricating oil pressure is built up before the engine starts to run.

During the winters of 1956–1957 and 1957–1958 actual service starting tests were conducted on all units operating on our proving grounds. In Illinois and Iowa there was an ample supply of cold weather.

All units, after having operated their usual 16-hr day, were parked in the open without any protection from wind and weather, and with sufficient space between units so that they did not shelter each other. Starting batteries as well as antifreeze were not removed from the units and neither heat nor chemical starting aids were used.

We were interested in whether or not the starting qualities of turbocharged engines differed from those of the naturally aspirated ones. We found that both types act alike, as can be seen from Fig. 47, which shows typical starting times of four of the units. These particular machines were selected as, prior to the initiation of the starting tests, each had operated for at least 1500 hr so that their mechanical condition was representative of contractor's equipment. The ambient temperatures preceding each start have been plotted to complete the information needed for evaluation of the tests.

It is gratifying that the engines can be started unaided in temperatures approaching −10 F but, with due respect to the life of the starting batteries, +5 F should probably be regarded as the approximate temperature below which the ether primer pump should be used judiciously.

Conclusion

We believe that the new models not only fulfill all of the requirements which had been established before the work was begun, but also that sufficient reserve strength was designed into these engines so that they are well able to cope with inevitable future demands for increased power.

Many years ago, Sir Harry Ricardo made this statement: "It is that heat which an engine fails to use that causes trouble, not that which it converts into useful work." This is the axiom which we have kept in mind throughout the job just completed, and which shall guide us in whatever development we may undertake in the future.

DISCUSSION

Supercharging Complicates Engine Development Testing

— K. J. Fleck
Caterpillar Tractor Co.

THIS IS a very straightforward, thorough report on a new and interesting development. The author has covered most questions that would normally arise, and should be commended on the interesting way he has described the problems and how they reached a satisfactory solution. In this discussion I will try to point out areas in which we concur with the thoughts expressed, and will try to ask a few questions which the author might discuss if he so chooses.

We recognize the value of single-cylinder testing. We began using single-cylinder engines during the development of our first diesel and have used them ever since for development of combustion and fuel-injection systems. We have experienced some difficulty in correlating single- and multicylinder performance, especially since supercharging has entered the picture. We find multicylinder engines necessary for completion of certain developments. For instance, we work out our final piston sizing on the multicylinder engine.

There still appears to be no substitute for field testing, and we too learn from our other divisions, especially the proving grounds and field testing groups.

There is not much said about the decision to use three valves. Was one of the purposes to permit a central location of the fuel spray nozzle? Also, was an optimum ratio of intake to exhaust port area established?

The theoretical swirl velocity analysis prompted some rough calculations. The indication is that if all the air were taken in during the intake stroke through the two valves at 2000 rpm in such a way as to create swirl only, the average velocity would be less than 200 fps. As only a small per cent of this air swirl seems to be necessary to mix the fuel and air, we might ask the author what air swirl is believed to exist in the actual combustion chamber. Fig. 11 indicates that the masks reverse the airflow and resulting swirl that could be had from the directional intake port. It appears that with full use of the directional port and the addition of more orifices, the masked valve could be eliminated. We feel that valve rotation is a good feature to have, fully realizing that many engines do operate satisfactorily without rotation.

The Allis-Chalmers engineers should also be commended in selecting a valve timing and other variables compatible with their system and in line with their objective and needs, rather than follow a general opinion blindly that high overlap was the only way to go. In the earthmoving industry the final engine design is sometimes a compromise for better low-speed performance. It is my belief that higher overlap might be useful only if chasing after horsepower at high speeds.

We agree with the philosophy of lubricating the fuel pumps with engine oil. With full-flow filtration, the engine oil is probably in better condition than a separate oil system would be without a filter.

Regarding the general operation and performance of these engines, the low fuel consumption is quite impressive. It is also noted that good performance is obtained without reaching exceedingly high cylinder pressures, according to today's standards. It might be of interest to know about what the "boost" or intake manifold pressure was on the supercharged engine. Also, due to the interest of the military services in multifuel engines, we might ask if these engines will operate on regular grade gasoline, and if so, what changes are required.

Swirl Found to be Equal With Directed Porting or Masked Valve

— L. D. Evans
International Harvester Co.

THE AUTHOR is to be commended on the fine presentation on this new engine. It is obviously the result of thousands of man hours of work and though the project

when put down on paper sounds simple and uncomplicated, it is never as easy as it is finally written. This is shown by the six year's development time.

There are a few points on which we wish to comment and also a number of questions which we would like answered.

First of all, it is stated that the directed port is not effective at the highest lift, whereas the masked valve is. We disagree with this view since our tests on directed porting indicate that swirl, even at high valve lift, can be equal to that obtained with masking and with no reduction in airflow. By means of offset counterbores, the initial swirl is good and the port direction carries on the swirl up to the highest lift. We agree that the masked valve is an easier method of producing swirl but we also feel that additional development will provide a satisfactory port eliminating the use of valve masking. As evidence of this, initially the M.A.N. "whisper" engine required shrouding but the latest configurations have eliminated the need for special masked valves, and many others get along without masking. Even though a satisfactory valve life is claimed by the author, the mechanism for controlling valve position, plus the additional valve cost, would seem to make it an attractive cost reduction item. During the development of the shrouded valves, were the valves rotated independently or together in obtaining the best position?

Again, relative to airflow we question whether Fig. 12 shows a condition of overswirl or of airflow reduction at the higher speeds with the 60-deg angle mark. We would also like to know the imep level of the curve shown.

In the original concept of this engine it was stated a normal valve overlap was decided on since piston temperatures could be controlled by oil cooling. However, later the author indicates that the forced oil cooling of the piston is not utilized and splash is the only device used for carrying oil to the piston. We are interested in knowing whether you can depend on splash cooling on crawler units which are operating continually on steep inclines in a process of oscillating up and down a slope doing bulldozing work. If splash is the only means of cooling are special oil flingers used? Our experience indicates that valves also suffer when low valve overlap is used at high output.

The firing pressures are not given in the text but on using the scale given, they appear to be in the neighborhood of 1350 psi at 165 bmep. This would be lower than we would expect. We wonder what the firing pressures were measured at 200 bmep? We also woud like to know what effect dropping to 14.5/1 compression ratio had on firing pressures and fuel consumption on the turbocharged engine?

Relative to the injection system we would like to know what injection pressures are measured and whether or not these were limiting pressures and determined the nozzle orifice size. Injection timing variance is also considered by many to be a necessity. Over the speeds shown, is there any reason why this is not being used? Were the performance curves drawn using best timing at each speed or were they at a compromise timing?

Relative to the mechanical aspects of the engine, the paper was not specific as to whether production crankshafts currently utilize rolled fillets. It is noted that a 40,000 psi iron is used in the crankcase. Is this checked on the crankcase and, if so, where? Also, what machining difficulties are encountered in using this hard iron? Is the production block cast upright or on its side?

The oil consumption for the naturally aspirated engine is shown and appears to be quite acceptable. We would be interested in knowing how the turbocharged engine oil consumption compares both in field operation and high constant load laboratory operation. On the performance curves we would like to know whether the horsepower and fuel rate shown are corrected and, if so, what correction was used for both the naturally aspirated and turbocharged engines.

As a matter of interest we would like to know the weight of the unit and the overall length from fan end to flywheel housing.

As a parting comment again it would appear development of this engine netted Allis-Chalmers a unit which has an enviable fuel consumption and features a good solid design.

Tangential Port Arrangement Found Advantageous by Mack

— B. Loeffler
Mack Trucks, Inc.

IT IS INDEED gratifying to note that Allis-Chalmers has also realized that an open-chamber engine offers a greater opportunity to the operators to save on fuel.

Not only does fuel economy attract the operator to the open-chamber engine, but also the many "hidden values" that go with it. These will manifest themselves only after thousands of miles or hours of operation.

Some of the hidden values are:

1. Improved lube oil consumption.
2. Longer piston and ring life.
3. Sustained power over long periods.
4. Quicker cold starts.
5. Longer exhaust valve life.
6. Reduced cylinder head cracking.
7. Reduced exhaust manifold cracking.

Type of Open Chambers

There are many ways to build an open chamber. It can be of the quiescent type with little or no air swirl. This would require a multiplicity of small spray holes with high pressure. It can also be of the higher air swirl type with less spray holes and moderate injection pressure. The latter type was chosen by the Allis-Chalmers Co.

To create the swirl there are the tangential and spiral port configurations and, then, there are the masked ports or masked valves. They all impart a swirling motion to the aspirated air inside the cylinder. Allis-Chalmers has chosen the masked valve idea while, for example, Mack has the tangential port.

We feel that the tangential port arrangement has several basic advantages over that of the masked valve.

1. A rotating intake valve keeps the valve seat clean by its sliding action, keeps the valve better sealed, and the seat doesn't look pockmarked.

2. The valve stem has no tendency to scuff. In fact, on a turobocharged engine it was found necessary to rotate the intake valve positively to prevent scuffing of the stem. As you know, the manifold pressure tends to remove the oil on the intake valve stem, aggravating the scuffing condition.

3. Another advantage is the flexibility with which the port velocity can be adjusted to the desired torque. To explain this, let us look at Fig. 12. On the left-hand side is shown the high torque at low engine speed, desired for a tractor engine such as the Allis-Chalmers. While a very desirable torque characteristic, it can't be used because the upper speed torque is very unstable. This is the result of the masked valve design. Had a tangential port been used, it would be possible to have the high torque at low speed and a uniform torque at the high speed.

The author mentioned that the A.E.C., Saurer, Fiat, and M.A.N. also have masked valves. M.A.N. switched to a spiral port. The latest Continental engine, also has the spiral intake port.

Two Intake and One Exhaust Valve

Inasmuch as Allis-Chalmers uses two intake valves, it seems superfluous to employ a mask. A slight change in the port shape would give at least the same results, so to speak, "free of charge." With two intake valves, the volumetric efficiency should be rather high; unfortunately, it is not recorded in the paper.

The combination of two intake and only one exhaust

valve intrigues me. Inasmuch as the engine breathes better, it seems to me that two exhaust valves would be better.

As Mr. Buchi pointed out, the least resistance offered by the valves results in improved specific fuel consumption. This is especially important when short timing events are used.

Valve Overlap

It is interesting that Allis-Chalmers has retained the same timing events for the turbocharged as well as for the naturally aspirated engine.

We, at Mack, came to the same conclusion after some six different camshafts were tried with overlaps up to 120 deg. Large overlaps gave a slight gain in specific fuel consumption at the top speeds, but ruined the part load, especially at lower speeds.

Such overlaps will do a creditable job on stationary engines, but this is quite different from automotive diesels that must operate over a wide speed and power range. Under those conditions, it is advisable to have an all around good economy.

Injection Advance

For slow-speed automotive diesels, there is not much need for automatic injection advance. For higher speeds, the automatic timing is beneficial because it reduces maximum pressures at low speed and smoke at high speed. Needless to say, an automatic advance will improve cold starting by many degrees.

Injection Pump

The Allis-Chalmers injection pump is engine lubricated. This was done, as explained, to eliminate dirt from entering the pump when re-oiling it. This is a good thought, but I doubt if it is the main reason for its use.

Some six years ago Mack adopted this method of pump lubrication because of certain fuels. These fuels, bordering on kerosene, leaked past the plungers into the cam box where it diluted the lube oil. Having little or no lubricating ability, it damaged roller, pin, and cams, and sometimes the ball bearings.

American Bosch devised the fuel return passage in the barrel as shown on Fig. 14. Lube oil splashing against the plungers sealed the leakage point which was nearly 100% perfect. In stopping the leakage, it was practical to engine lubricate the moving parts.

Of interest is the pump drive. Being exposed, it allows the tinkerer to manipulate the pump timing to suit his fancy. It is not necessary to enumerate the trouble that can result from such a drive arrangement.

We have found it necessary to make the connection between pump and drive "tinkerproof." It is done by completely enclosing the drive and pump timing adjustment and incorporating tamperproof seals for the drive, pump rack and governor.

Fuel Consumption

As expected from a 4-stroke open-chamber engine, the specific fuel consumption, as shown in this paper, is good.

What was the reason for reducing the ratio from 15.5 to 14.5? What is the compression pressure with the reduced ratio over the speed range and at cranking speed?

Author's Closure To Discussion

WE WISH to express our thanks to the discussers for the interest which they have shown in our presentation. Quite a few questions were asked and we will attempt to answer them in the order in which they appear. However, since several discussers, in writing as well as from the floor, requested rather similar items of information, it seems advisable to provide these, in the form of graphs, following our individual replies.

Our experience agrees with Mr. Fleck's, namely, that certain problems of the turbocharged engine cannot be solved on experimental single-cylinder units. As he points out, piston sizing has to be done in multicylinder engines and not only on the dynamometer but also on the proving ground.

The decision to employ a central nozzle location preceded that of using three valves. A single exhaust valve, rather than two, was used to avoid the sensitive cooling water passage which is required between double valve pockets. Valve area ratio was not experimented with, as the largest valve diameters which still permit safe casting sections between valve seat and nozzle were used.

Mr. Fleck's calculation concerning average air inlet velocity is arithmetically correct but based on an impossible assumption — namely, that mean inlet velocity and tangential velocity are the same. We must remember that the air velocity component, which is parallel with the cylinder axis, determines the degree of filling. It can be shown mathematically that this vertical component, to give acceptable volumetric efficiencies, must be at least as great as the tangential component. This is the basis for the assumption made in the paper, namely, that actual velocity must be at least 1.41 times the required initial swirl velocity. Near upper dead center and at 1800 engine rpm, the actual swirl velocity in the chamber is very close to four times crankshaft speed. Fig. 12 shows that this swirl velocity is actually required to obtain the high combustion efficiency achieved. Lower swirl velocities, resulting from too small or too large a mask position in combination with correct mask angle, reduce the torque which is produced by a constant fuel quantity.

Fig. 11 is an artist's conception and was used in the paper only to show the cross-section of the engine and not the actual airflow pattern. The airflow in the engine is not "reversed" as the masks are placed up-wind, that is toward the inlet manifold. In plan view, their locations coincide fairly well with the transverse centerline of the port.

With respect to valve timing, both Messrs. Fleck and Loeffler, have mentioned a very sound reason for keeping valve overlap quite low: Low-speed torque backup, that is high volumetric efficiency at low engine speed, is just as important as maximum horsepower when engines are intended for automotive applications using mechanical transmissions. A large overlap would obviously not contribute to satisfying this requirement. Another, and apparently not fully appreciated, benefit of reasonably small overlap consists in its beneficial influence upon turbocharger response time.

Mr. Evans is quite right when he states that ours is rather a costly valve arrangement. It is not possible, however, to judge an important design detail on a cost basis only. We feel that the added expense is warranted because of the greatly improved performance; and, as already stated, because proper swirl velocity control over the whole operating speed range eliminates the need for a variable timing device.

Valves were rotated individually and different mask angles for each valve were tried, but proper duct design resulted in optimum mask angle and position being the same for both valves.

The dip in the torque curves in Fig. 12 definitely represents overswirling as there was no measurable change in airflow. The indicated mean effective pressure for the 45-deg mask angle curve at 100-deg position is approximately 140 psi at all speeds. So far, forced oil cooling of the pistons is not used for production ratings. We have not noticed any difference in piston temperature between level and inclined operation, probably due to the fact that the oil throw-off speed from the crank pin is sufficiently high so that the effect of gravity is not significant.

Regretably, the fuel pressure scale was omitted from Fig. 46. It is 1000 psi per line on the Y axis and the "pips" on the X axis are 5-deg markers. Thus, the maximum pressure, at 1800 rpm and maximum fuel quantity, is of the order of 7000 psi, which is so far below the limiting pressure that we did not have to make any compromise with respect to either nozzle hole size or to any other parameter of the injection system.

We have no reliable data on firing pressures with the original 15.6/1 compression ratio but we are agreed that the final 14.5/1 ratio makes the engine sound "sweeter." There is no measurable difference in specific fuel consumption between the two ratios, either in the naturally aspirated or in the turbocharged engine.

As mentioned under "Data Presentation," all performance data are observed without any correction. Fixed timing was used throughout.

All crankcase fillets are rolled. The cylinder block is cast lying on the push rod side. A 1.2 in. diameter test bar is cast with each block and checked for tensile strength and Brinell hardness. The latter test is then compared with Brinell tests on pan rail and top deck.

We have no complaints from the production department about machining difficulties, probably because foundry control is very rigid and castings are carefully stress-relieved.

On constant load dynamometer operation the lubricating oil consumption of the turbocharged engine is some 50% higher than that of the naturally aspirated one. In other words, oil consumption per bhp-hr worked out about the same. In field service in earthmoving machinery we expect approximately 0.7 qt per 8 hr of operation.

The standard "21000" commercial engine, with starter, generator, fan, and air cleaner, weighs 3215 lb and the "16000" is some 130 lb lighter. Either engine measures 63½ in. fan to flywheel.

We fully agree with Mr. Loeffler concerning the "hidden values" of this type of engine. Only now, as large numbers of these engines are accumulating long service hours, do we begin to fully appreciate the items listed by Mr. Loeffler.

We might also agree with his first two points concerning rotating versus keyed valves if we were trying to push these engines. For ratings so far used commercially we have had no reason to change our original thinking.

We disagree with the third point as we fail to see anything "flexible" about intake velocity when it is imposed upon the air by the shape of a cast-iron duct. It is certainly easier to change mask position, and even mask angle, than to make new port cores and, from them, new cylinder head castings. We have stated previously that a single intake valve lends itself better to directed port design than do dual valves. Normal cylinder head height, the usual cylinder distance, and the necessary stud spacing pretty well defined the shape of the port. In the case of the valve nearest the intake manifold, the bottom of the port can be shaped properly but the air lacks guidance near the far side of the valve. Proper directional porting of the farther valve is extremely difficult to achieve as the bottom of the duct must clear the nearer valve and, thus, cannot be made flat enough. In plan view, lack of space between bores prevents correct tangential placement of the valve end of the port. Thus, dual intake valves require some design assistance if they are to produce optimum

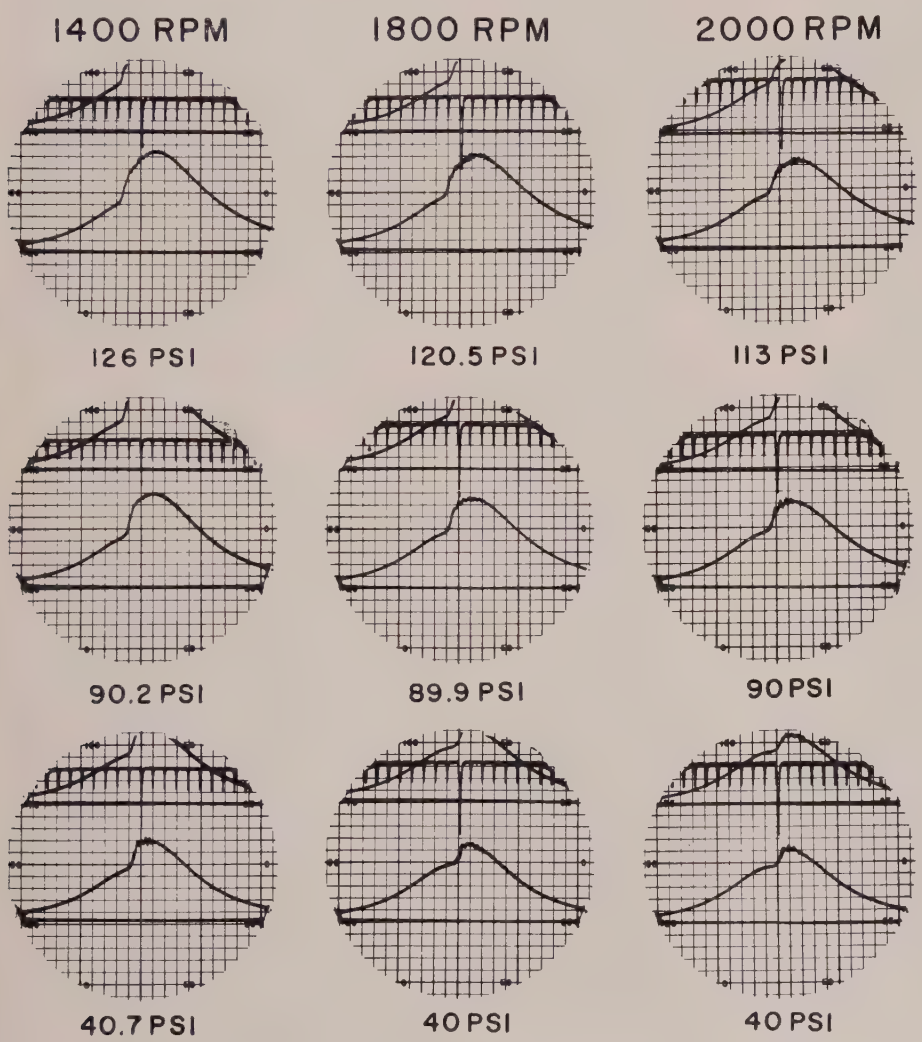


Fig. A — "16000" cylinder pressures, fixed timing (Y axis: 1 line = 100 psi, X axis: 1 mark = 5 deg crankshaft rotation)

swirl rates. This, however, is a small penalty to pay for higher volumetric efficiency, lower valve weight, lighter spring loads, and improved cylinder head geometry.

Several of the discussers asked questions concerning cylinder pressures at load conditions other than those shown in Fig. 45, and they also inquired about intake manifold and differential pressures. It would hardly be possible to give satisfactory written answers to all these questions and for this reason four additional figures have been prepared which, we hope, will provide clarification.

Fig. A shows that even in the naturally aspirated engine and, at that, at light loads, the transition between compression and firing curves remains smooth and the rate of pressure rise extremely gentle.

In connection with Fig. B it must be noted that in frames 31, 33, 20, 21, 13, and 15 the longest "pip" is not at top dead center but at 5–6 deg thereafter, a condition due to the fact that, at the time when these diagrams were taken, we did not yet realize the need for continuously checking the adjustment of the synchromarker drive.

The relationship between peak pressure, bmep, and engine speed is plotted in Fig. C, based on mean values obtained from several pressure diagrams for each operating condition. We have found such averaging to be necessary as consecutive photographs of pressure traces, at the same condition, show noticeable variations when a normal exposure time of 1/10 sec is used. The reason for this becomes clear when it is realized that, at 1800 rpm and 1/10 exposure time, the photographed trace represents only one and one-half engine cycles. As the oscilloscope is triggered at the beginning of the suction stroke, the trace, from about half-way up the compression stroke onwards, is produced by only one single event. Thus, it is necessary to establish the true mean condition before reliable information can be obtained.

Figs. B, C, and D pertain to the standard "21000" for which the performance is given in Fig. 41. Fig. A shows pressure traces from the standard "16000" with the same settings which were used in obtaining the data on Fig. 39. As in all other graphs, timing remained constant throughout the tests at the value established for production engines.

In closing, we would like to state that we have no doubt that further development work on these engines will be rewarding. It would be rather presumptuous if we were to assume that, after just a few years of development, we had already realized the full potential of these models. As a matter of fact, since the paper was written (Spring, 1959) noticeable progress has been made, especially in the area of higher sustained outputs.

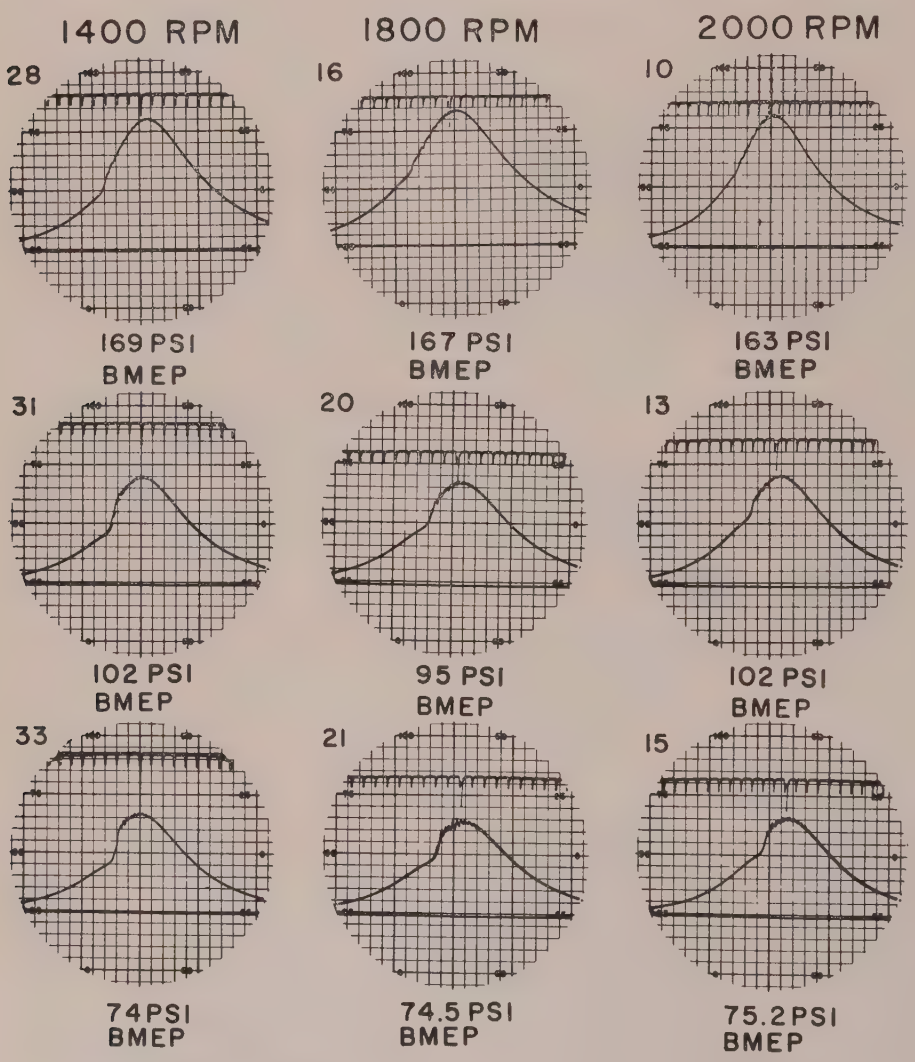


Fig. B — "21000" cylinder pressures (*Y* axis: 1 line = 100 psi, *X* axis: 5 deg markers)

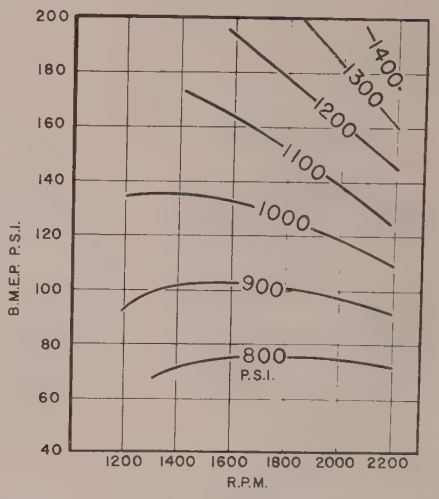


Fig. C — "21000" maximum pressure map

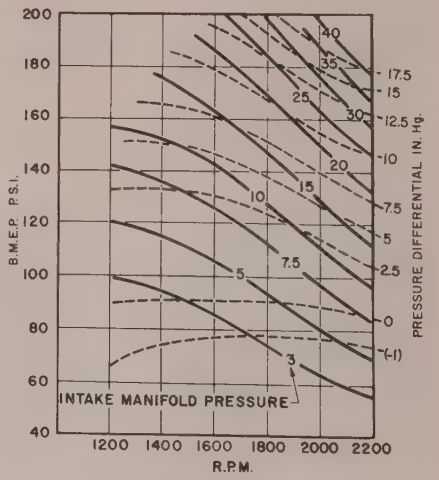


Fig. D — "21000" intake manifold pressure and pressure differential map

New ELASTOMERS Break Old Barriers

Ralph P. Schmuckal
Ford Motor Co.

THE AUTOMOTIVE industry has not yet taken full advantage of all the superior and peculiar properties of the new elastomers. To accomplish this, components must be designed around the material in contrast to the more normal procedure of fitting a material to an established design.

Examples of progress are: colored rubbers that withstand weather indefinitely; elastomers that seal the most deteriorating hydraulic and lubricating fluids at temperatures in excess of those practical from the standpoint of the stability of the fluids themselves and the effects of heat upon metal components; the fact that the long-sought 100,000-mile tire is imminent.

This paper discusses the properties and potentials of some of the new elastomers, with some possible applications.*

NEW ELASTOMERS and modifications of old ones continually are being developed in the fast-moving and exciting field of nonmetallic materials. The automotive engineer adapting these elastomers to his product is exposed to myriads of claims and counterclaims, with emphasis invariably being placed on the advantages or strong points of a given material. Not too much is heard about deficiencies. Materials are made to look good. Some of them are truly remarkable. Others, due to cost or inherent weaknesses in the polymer itself, find their use rather limited in today's keen competition for a place in the automotive industry.

Today's automobile, with over 500 rubber parts and potentially hundreds more, presents a real challenge to the designer and materials engineer to obtain "the most for the least" or in practical terms, to get adequate performance at a minimum of cost. The principal purpose of this paper is to present to the automotive engineer not fully versed in elastomers and their properties a condensed package of information on each of the newer materials—their molecular structure, their advantages and shortcomings, some current and potential applications, and the all-important price picture. The new elastomers are compared to existing materials and fitted into their proper niche in the elastomeric scheme of things.

It will be seen that many of the old barriers prohibiting the use of elastomers have been broken. Freedom of design is open to the imaginative designer with elastomers available today ready to perform jobs merely dreamed of just a few years ago. Colored rubbers withstand weather indefinitely. Elastomers seal the most deteriorating hydraulic and lubricating fluids at temperatures in excess of those practical from the standpoint of the stability of the fluids themselves and the effects of heat upon metal components. The long-sought 100,000 mile tire is imminent. And the field of elastomers is yet in its infancy.

The following discussions cover the properties and potential of the new elastomers in some detail.

Isoprene Rubber

The shortage of natural Hevea rubber during World War II, particularly for tire use (as nearly everyone clearly remembers), led to intensified efforts to develop suitable substitutes. GR-S did an outstanding job of filling the requirements of a polymer for passenger-car tires and most other mechanical rubber parts formerly made of Hevea rubber. For large tires, however, GR-S was not entirely satisfactory because of excessive heat buildup. After many and long explorations of synthetics for large truck and military applications, it was decided by the rubber industry that the natural rubber molecule must be built up synthetically in order to obtain the necessary properties of strength, elongation, abrasion resistance, processability, and the all-important low hysteresis property so necessary for such tires. Isoprene, the natural rubber monomer, became the building block for the new synthetic "natural" rubber. The work, carried on independently and simultaneously by all the major rubber companies, was highly successful and gave us a cis-

* Paper presented at SAE Summer Meeting, Atlantic City, June 19, 1959.

polyisoprene with physical properties and polymer structure very similar to those of Hevea rubber (Fig. 1).

I. R. (isoprene rubber)[1,2,3] tire tread stocks show room-temperature tensile strengths around 4000 psi with elongations in excess of 500%. Graphical comparisons for these two properties of I. R. and Hevea with a comparable SBR (formerly known as GR-S) over a wide temperature range are shown in Figs. 2 and 3. It is seen that I. R. and Hevea rubber have similar retention of both tensile and elongation up to 350 F. At intermediate temperatures (150–250 F), both are superior to SBR in tensile and elongation. At extremely high temperatures (350 F) all three polymers have about the same tensile, but SBR loses most of its elongation.

After heat aging in the range of 250–300 F, I. R. retains its physical properties nearly as well as Hevea rubber.

As stated earlier, I. R. was developed for its low hysteretic properties without sacrificing the other desirable properties of true natural rubber. Fig. 4 shows graphically a comparison of heat buildup for I. R., Hevea, and SBR on a sheer flexometer. Hysteresis of I. R. is very nearly that of Hevea rubber and less than half that of SBR.

Tire tests have confirmed laboratory data showing that cis-polyisoprene is 95–100% as good as natural rubber for wear and for heat buildup during tire operation. It also is as good at low temperatures, giving comparable traction on ice and snow. I. R. treads have less tendency towards cracking than does natural rubber. It has been successfully tried in engine mounts and bushings for automobiles.

Commercial prospects of isoprene rubber appear very good. Widespread use has awaited the lower cost of isoprene. Indications are that in the future the cost of cis-polyisoprene will be competitive with natural rubber but probably never as low as that of SBR. Another important contribution of the synthetically produced polyisoprene to the national economy is that by its very existence it will tend to hold the price of natural rubber down and stable. Furthermore, large stockpiles of natural rubber for national emergencies no longer will be required.

For practical purposes, isoprene rubber may be considered identical to Hevea rubber. However, I. R. has fewer impurities than the natural polymer and, therefore, is more uniform in quality.

Several companies are building plants for commercial production of isoprene rubber. One plant announced an annual capacity of 30,000 tons.

In full production, the cost of I. R. is expected to be no higher than current costs of natural rubber.

A number of typical examples of isoprene rubber are known by trade names of Coral Rubber, Ameripol SN, Natsyn Rubber, and Shell Polyisoprene.

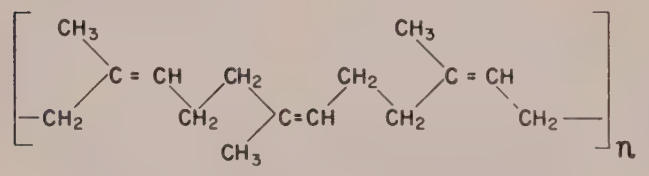


Fig. 1 — Cis-polyisoprene (isoprene rubber)

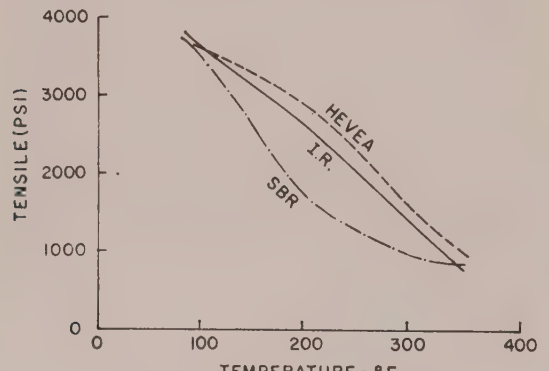


Fig. 2 — Tensile versus temperature, I.R., Hevea, and SBR compounds

Butyl Rubbers

Conventional butyl rubber has become a widely used general-purpose (nonoil-resisting) rubber. Chemically it is a copolymer of isobutylene and small amounts of isoprene, with a molecular structure as shown in Fig. 5. Variations in butyl polymers exist, depending principally upon the mole ratio of isoprene to isobutylene. The more isoprene that is used, the greater is the chemical unsaturation. Present commercial butyl polymers range in percent mole unsaturation from 0.9 to 2.2. The lowest unsaturation is used where maximum ozone and chemical resistance are required, such as in wire and cable, pipe, tank linings, and the like. Care must be taken in the processing of this stock. The highest unsaturation is fastest curing, easiest processing, and is used primarily where heat resistance is required, such as is curing bags and bladders, carcass and tread base compounds for tires, and the like. Grades between the two mentioned represent compromises in such physical properties and are widely used or finding use in such items as inner tubes, weatherstrips, suspension bumpers, boots, convertible tops, portable swimming pools, rainwear, hydraulic seals.

Summarizing a few of the desirable properties of butyl, it has:

1. Good heat resistance.
2. Good ozone and weather resistance.
3. Low permeability to gas.
4. Good tear resistance.
5. Good flex resistance.
6. Good resistance to synthetic fluids, such as ester and glycol types which are very polar in nature.
7. Good colorability (in molded items).
8. High damping characteristics.

Summarizing a few of the shortcomings of conventional butyl, it has:

1. Only fair compression set characteristics.
2. Low resilience at normal and low temperatures.
3. Poor resistance to mineral oils and fuels.
4. Incompatibility with other elastomers.
5. Somewhat difficult processing characteristics.
6. Only fair low-temperature properties.

[1] "Cis-Polyisoprenes and Polybutadienes — Properties and Uses," by Glen Alliger. *Rubber World*, Vol. 139, March, 1959, p. 891.

[2] "Properties and Applications of Coral Rubber — Cis, 1, 4 Polyisoprene," by G. Alliger, J. M. Willis, W. A. Smith, and J. J. Allen. *Rubber World*, Vol. 134, July, 1956, pp. 549–559.

[3] "Military and Economic Impact of Coral Rubber," by F. W. Stavely. Firestone, March 28, 1958.

Much published information is available on conventional butyl rubber—its chemistry, properties, and applications—making it a widely known material. Not so well-known are some of the more recent modifications of butyl which overcome some of the above mentioned shortcomings of conventional butyl.

Halogenated Butyl

Halogens, such as bromine and chlorine, added along the butyl chain make important improvements in the properties of butyl. Polymers with 1.0–3.5% bromine have significantly increased cure rates and are compatible with other elastomers.[4] This modified butyl actually can be blended with natural rubber and synthetics to give desirable properties of each material. Conventional butyl cannot be blended in any such manner; in fact, processing butyl in proximity to other more unsaturated type materials like SBR, Buna-N, and Neoprene is difficult, due to the fact that small amounts of such polymers rob the system of vulcanizing agent and prevent proper curing of the butyl.

In mixtures with natural rubber, bromobutyl reduces the tensile strength of natural rubber but not below a respectable value. Also, the gas diffusion rate through natural rubber or SBR is greatly reduced by bromobutyl, decreasing linearly with increasing bromobutyl ratios up to a 50-50 mixture. Likewise, ozone resistance of SBR and natural rubber are appreciably improved by blending with bromobutyl.

Brominated butyl exhibits adhesion to metals and other rubbers superior to that of conventional butyl.

Because of its adhesive qualities, low gas permeability, and compatiblity with other elastomers, bromobutyl is used in the tire industry in inner liner compounds of SBR tubeless tires. It retains air like butyl rubber and adheres to the carcass like SBR or natural rubber. Bromobutyl rubber is used as a tie ply to retread an SBR tire with butyl tread. The polar halogen atom apparently accounts for its compatibility.

Chlorobutyl[5] rubber, similar chemically and physically to bromobutyl, is a new polymer consisting of the basic butyl polymer with 1.1–1.3% chlorine (by weight) attached along the chain. It is uncertain at this time whether it attaches to the double bond or replaces hydrogen atoms of the CH_3 or CH_2 groups. Today it is being produced in pilot plant quantities only, and has not been fully evaluated for specific uses. Like bromobutyl, it is compatible with other elastomers, has good adhesive qualities and air holding properties. With a sulfur cure system, tensile strengths of 2800 psi and elongations of 300–400% are obtained.

With chlorobutyl, an unusually large number of cure systems are available, each imparting somewhat different properties to the vulcanizate. The wide selection of cure systems and blending possibilities with other elastomers make it possible to produce compounds covering an extremely wide range of physical properties and to accentuate certain properties in a given compound. Some chlorobutyls have exceptional ozone resistance, others unusually high heat resistance, flexing performance, tear strength, compression set, or abrasion resistance. Laboratory tests indicate certain chlorobutyl compounds and blends of chlorobutyl and butyl to have abrasion resistance twice that of regular butyl tire tread stock. It appears, therefore, to have great potential as a tire material. A comprehensive study of this potential is in progress. Another interesting feature of chlorobutyl is that oil-resisting materials like Buna-N and Neoprene can be blended with it to improve its oil resistance without sacrificing too much the desirable qualities of butyl.

Chlorobutyl's low-temperature and damping properties are comparable to those of butyl.

It is possible that chlorobutyl, in addition to being a specialty material, will become a general purpose elastomer. In addition to the uses already cited, possible applications are molded items requiring good high-temperature service, such as gaskets, couplings, ring seals, brake boots, vibration dampers, spark plug boots, hoses, and similar applications where only a small degree of oil resistance is required.

The price of the chlorobutyl polymer with a specific gravity of 0.92, is about $0.30 per pound, about

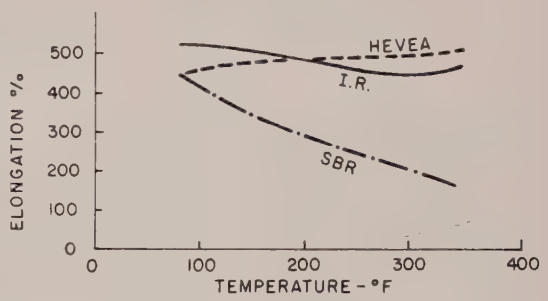


Fig. 3 — Elongation versus temperature, I.R., Hevea, and SBR compounds

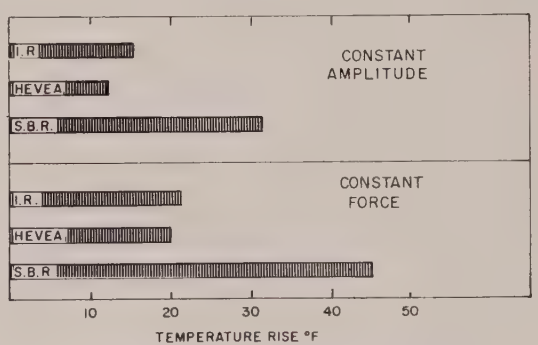


Fig. 4 — Heat buildup of I.R. as measured on shear flexometer (re: *Rubber World*, July, 1956)

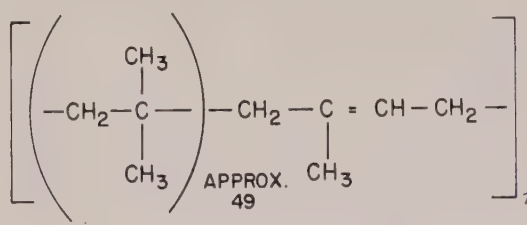


Fig. 5 — Butyl

[4] "Meetings and Reports." *Rubber World*, Vol. 139, March, 1959, p. 889.
[5] "MD-551 — A New Synthetic Elastomer." Technical Bulletin No. 9. Enjay Co., Inc., New York.

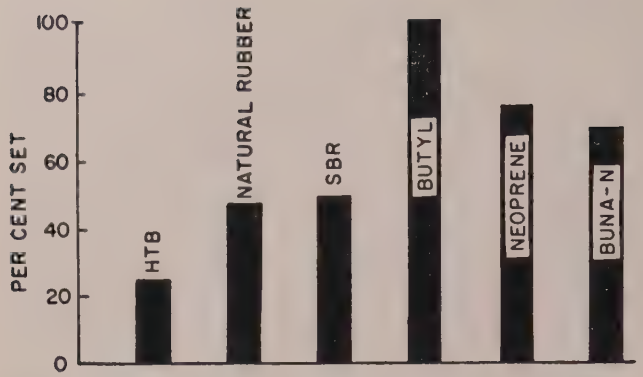


Fig. 6 — Compression set of HTB versus other typical compounds at 300 F and 30% compression deflection

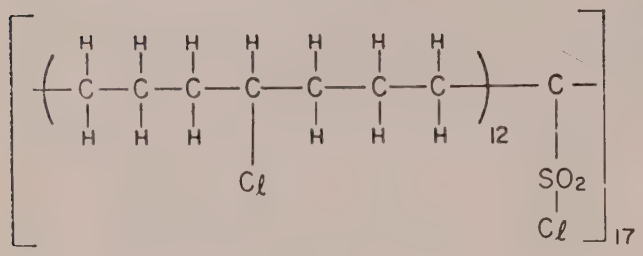


Fig. 7 — Chlorosulfonated polyethylene

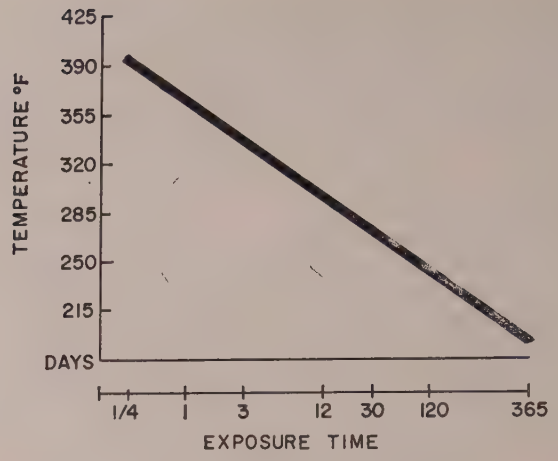


Fig. 8 — Heat aging of chlorosulfonated polyethylene, time to drop elongation to 100% (E. I. du Pont)

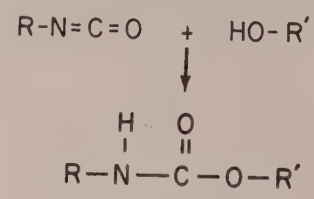


Fig. 9 — Urethane (where R and R' are hydrocarbon radicals)

$0.06 per pound above that of the standard butyl polymer.

High-Temperature Butyl

High-temperature butyl (HBT) sometimes is referred to as a "resin-cured" butyl. It is not a new polymer, but the product of a development in curing systems of the butyl polymer. Phenol resins act as the vulcanizing agents. This cure imparts unusual thermal stability, ozone resistance, and compression set to the butyl polymer.

The top temperature for continuous exposure to dry heat has been raised to 350 F. Periodic exposure to peak temperatures of 400 or even 450 F can be tolerated. Such performance rates it well above such elastomers as natural rubber, SBR, Buna-N, conventional butyl, and neoprene. Fig. 6 shows this heat resistance in a different manner, by comparing its compression set properties, at 300 F and 30% deflection, with other well-known polymers.[6] The graph indicates that at 300 F the compression set of high-temperature butyl is only about one-quarter that of standard butyl. Other similar data not shown here indicate that high-temperature butyl at 300 F is as good as conventional butyl at room temperature.

Resin curing does not significantly alter the rather poor oil and gasoline resistance of butyl. Nevertheless, this material should find use as seals for certain hydraulic fluids, possibly engine mounts and bumpers, elastomeric backing materials to maintain preload, gasket materials, and mechanical rubber parts in the engine compartment of an automobile where electrical properties, heat resistance, ozone resistance, and possibly abrasion resistance are requirements, but where the material does not come into continuous contact with mineral oils or fuels. Spark-plug boots and brake cups are additional possibilities.

The price of high-temperature butyl is similar to that of regular butyl.

Chlorosulfonated Polyethylene

Chlorosulfonated polyethylene has been available for several years; nevertheless, a discussion of it is included in this paper since it still is considered by the automotive industry as a specialty material that slowly but continually is finding wider usage.

Chlorosulfonated polyethylene is made by reacting polyethylene plastic with chlorine and sulfur dioxide. The resulting molecular structure is shown in Fig. 7. This material, unlike thermoplastic polyethylene, is vulcanizable with properties of a true elastomer.

This material is noted for its resistance to heat, ozone, weather, abrasion, and chemicals. Color stability is excellent.

Its heat resistance may best be shown by a plot[7] of temperature versus exposure time to reduce the elongation to 100% (Fig. 8). At 100% elongation the material is considered still sufficiently elastomeric for most service applications. It is seen from the graph, for example, that the service life at 350 F is over one day; at 300 F, about 9 days; at 250 F, about 3 months. This is under continuous, unprotected exposure to hot air. Such resistance to heat is considerably better than that obtained with neoprene which also is considered to be a good heat-resisting material. To make a direct comparison, neoprene, aged 9 days at 300 F, would have less than 20% stretch remaining.

The ozone resistance of this material is excep-

[6] "High Temperature Butyl," by G. S. Buettner and C. R. McGill. U. S. Rubber Co.

[7] "Facts about Hypalon." E. I. du Pont de Nemours & Co., August, 1955.

tional. In fact, it is completely unaffected by ozone, even at elevated temperatures and extremely high ozone concentrations.

Its chemical resistance may be summarized as follows:

1. Outstanding resistance to acids and strong oxidizing agents such as sulfuric acid and hydrogen peroxide.

2. Good resistance to nonoxidizing chemicals, such as ethylene glycol, alkalies, and the like.

3. Resistance to mineral oils nearly equal to that of neoprene. It is unsuitable for direct contact with gasolines and aromatic solvents.

Weather resistance is very good, even for light-colored materials. Exposure in Florida for 2½ years resulted in no noticeable color change or significant loss in physical properties.

Regarding its low-temperature properties, practical vulcanizates stiffen at somewhat higher temperatures than more conventional elastomers. The 10,000 psi (Clash-Berg) modulus occurs in the neighborhood of 0 F. However, by special compounding and at some sacrifice of other desirable properties, this degree of stiffening may be reduced to temperatures as low as – 60 F. An important point to be noted is its extremely low brittle point, even for compounds that stiffen at relatively high temperatures. Brittleness usually is encountered in the range of – 60 to – 100 F.

It may be well to point out that many rubber fabricators have resisted the extensive use of chlorosulfonated polyethylene because of difficulties in processing this material in their plants. Possibly this fact contributed in part to the development of a new chlorosulfonated polyethylene which is planned to be introduced commercially in the near future. Its primary advantages over the currently used material are its oil resistance, which is comparable to that of neoprene, its greater strength, and its much improved processability.

With these properties, chlorosulfonated polyethylene should find use in the automotive industry. Its ozone, heat, and oil resistance, together with good electrical properties and abrasion resistance, make it a desirable material for spark-plug boots, ignition wires, white sidewalls of tires, convertible topping, weatherstrip coatings, coatings for urethane foam headliners, and truck tarpaulins. The present trend to colored material also may bring about a demand for colored windshield wipers and colored window weatherstrips, possibly of the material, in the future.

The price of the polymer is about $0.70 per pound. The price of a typical compound, with a specific gravity of 1.4, is $0.45 to $0.50 per pound.

Chlorosulfonated polyethylene is marketed under the trade name, Hypalon.

Urethane Rubber

Solid urethanes constitute a large family of materials produced basically by combining diisocyanates with polyesters or polyethers, followed by curing or crosslinking with peroxides, amines, glycols, heat, and so forth. The chemistry is similar to that used in producing urethane foams, but without the blowing agent that produces the cellular structure.

The complete chemistry involved in urethane production is very complex. The basic, essential reaction to form a urethane is the reaction of an isocyanate group of one material with the hydroxyl group of another (Fig. 9). Cross-linking or polymerizing this urethane into large, complex molecules completes the process of forming the urethane elastomer.

Many varieties of urethanes with a wide range of physical properties are produced. Adiprene-L and Vulcollan[8] are liquid, castable types. Adiprene-C is a gum-type polymer, processable on conventional rubber processing equipment. Unlike the liquid types, Adiprene-C can be reinforced with fillers like carbon black. Similar or modifications of these urethanes are commercially available under the trade names Genthane, Vibrathane, Estane, Daykollan, and Chemigum SL.

For the automotive engineer, this material can be described best in terms of its physical properties, as follows:

Hardness—Urethanes are fairly hard materials, usually falling in the 60-98 Shore A durometer hardness range, with 80–90 durometer most common.

Tensile—Tensile strengths of 4500–7500 psi are not uncommon. In general, the harder the stock, the higher the tensile.

Elongation—Per cent elongation on most stocks is exceptionally high—400–750%. One available stock of 80 durometer hardness and 7280-psi tensile elongates 640%. Its elastic properties are exemplified by the fact that its elongation set, after this high stretch, is only 20%.

Tear—Tear strengths also are remarkably high and, in general, increase as hardness values increase—ranging from about 250–700 lb/in. Compare this with about 300 lb/in. for a high tear strength neoprene.

Abrasion Resistance—Laboratory accelerated wear tests show from three to ten times the wear of other synthetic rubbers or natural rubber.

Ozone Resistance—Ozone resistance is excellent. Slight surface cracking after 16 hr in the abnormally high ozone concentration of 10,000 pphm is typical. Outdoor weather conditions will not degrade urethane, but do tend to discolor it.

Heat Resistance—Urethane is not materially affected by dry air temperatures up to 212 F. About 250 F should be considered the upper limit of serviceability.

Low-Temperature Properties—Typical urethanes stiffen as the temperature is reduced, but show little tendency to crystallize. Although rather stiff at sub-zero temperatures, urethane does not become brittle down to – 80 F.

Resiliency—Urethane resilience usually falls in the range from 50% to a maximum of 80%, which in general is higher than that of butyl but lower than that of most of the other conventional elastomers.

Oil Resistance—To summarize urethane's oil resistance in mineral oils, it is better than neoprene but inferior to Buna-N, provided temperatures do not exceed 200 F. Urethane rubber gradually deteriorates in 250 F mineral oil and is completely unserviceable at 300 F. Urethane is not recommended for use with most synthetic hydraulic fluids, hot water, or steam.

The high modulus and high elongation of urethane, together with appropriate hysteresis, pro-

[8] "Vulkollan Urethane Rubber." Mobay Chemical Co.

vide excellent energy absorbing power and vibration dampening properties. Its toughness can be illustrated by its early application as a mallet head. In this item it is hard enough to drive a ten-penny nail into a solid oak beam without damage to itself. Industrially it is finding wide usage as shock-absorber pads in and under heavy machinery, as air hammer handles to absorb shock, suspension bearings on heavy equipment, and flexible couplings. Its strength and flexibility permit it to transmit high loads with high misalignment.

Its abrasion and tear resistance early suggested urethane's use as a tire material. Hard rubber tires are now commercial and automotive tires are experimental. This material may produce the awaited 100,000-mile tire.

Where temperatures are not excessive, urethane makes a good oil seal. But in today's engines and transmissions, urethane is not being used because of the temperature encountered.

Extremely hard urethane materials, internally lubricated, are being used abroad as self-lubricating bearing materials and are being evaluated in this country for that purpose.

The urethane rubbers appear to have tremendous potential in the mechanical rubber goods field because of their exceptional physical properties.

Table 1 — Effect on Typical PA Seal Stock of Aging One Week in Various Media

	Tensile Change, %	Elongation Change, %	Volume Change, %
Air at 300 F	−21	−14	
350 F	−85	−48	
ASTM Oil No. 3 at 300 F	− 2	+14	+17
350	− 4	+14	+17
ASTM Oil No. 4 at 300 F	− 1	+43	− 2
350 F	+43	−52	− 8

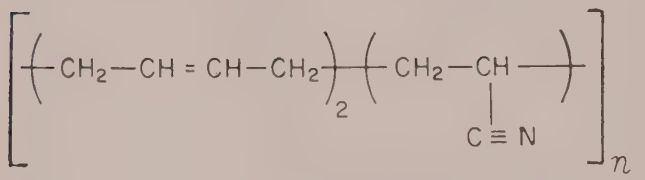


Fig. 10 — Buna-N (NBR)

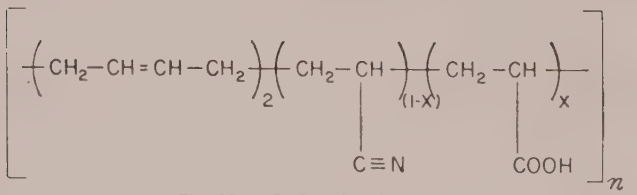


Fig. 11 — Carboxylated Buna-N

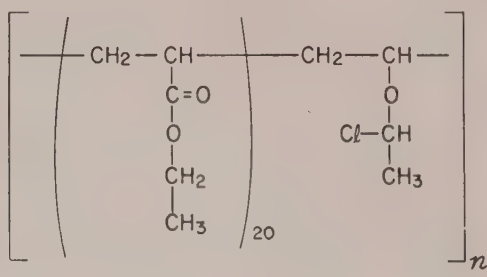


Fig. 12 — Polyacrylic rubber

However, they are not yet used extensively in the automotive industry for such applications, partly because of cost. This material, being so different from conventional elastomers, should not be considered as a direct replacement for currently used materials. To take full advantage of its superior properties, the component must be designed around the material. In highly loaded mechanical rubber parts, far less volume of urethane can be made to do the same job. The resultant savings in rubber volume in many cases would offset the higher cost per pound and give additional bonus features in reduced weight and space requirements. If low-temperature stiffening does not prove to be an insurmountable problem, urethane may some day make a compact, "super" engine mount. However, converting to urethane in such an application is not an easy step and certainly considerable developmental work would be required.

Urethane is also being considered for use as ball-joint liners, gears, and for propeller shaft and universal joint couplings. Urethane also has great potential as a coating material on exposed metallic, wooden, and rubber parts, and over flexible urethane-foam upholstery.

The price of high-grade urethane elastomers, with a specific gravity of about 1.1, is about $1.15 per pound.

Nitrile Rubbers

Buna-N, the most common nitrile rubber, is a copolymer produced from building blocks of butadiene and acrylonitrile. The molecular configuration appears in Fig. 10. Although conventional Buna-N is not a new polymer, it may be well briefly to describe some of its properties and uses as background material before discussing some of the more recent modifications of nitrile rubber.

Butadiene and acrylonitrile can be polymerized in various mole ratios. The higher the acrylonitrile content, the greater the resistance to petroleum oils, fuels, solvents, and heat. The higher the butadiene content, the greater the resiliency and low temperature flexibility. Buna-N polymers, therefore, can be "tailor-made" for specific end uses.

Buna-N withstands dry heat up to about 250 F and most petroleum-based oils up to about 300 F. Abrasion resistance is good. Compression set properties are excellent. Properly compounded, it has no difficulty in meeting a − 40 F cold flex test. These properties are responsible for its extremely wide usage in engines, automatic transmissions, and power steering devices, as reciprocating and rotating shaft seals, gaskets, tubing, coated fabric diaphragms, and the like.

For dynamically functioning rubber parts, Buna-N does not have the flex resistance of neoprene or the general purpose elastomers. Ester-type hydraulic fluids, extreme-pressure lubricants, and other sulfur-bearing fluids severely degrade Buna-N, particularly at elevated temperatures of 250 F or above. Some of these deficiencies of Buna-N led to the development of polymer modifications and blends as follows.

Carboxylated Buna-N

This polymer is a medium high acrylonitrile copolymer which has been modified to include carboxylic groups in the polymer chain, as shown in

Fig. 11. The carboxylic group imparts desirable properties to the material, such as outstanding abrasion resistance, better retention of physical properties at elevated temperatures, and somewhat improved ozone resistance. The abrasion resistance of carboxylated Buna-N, on a laboratory NBS abrasion test, is three to six times as great as that of conventional Buna-N. Ozone resistance is significantly improved, but not of the order of that obtained with neoprene or butyl, unless compounded with waxes for ozone protection. Immersion in oil further improves ozone resistance. Resistance to the aging effects of dry heat and oils is not materially altered.

Improved abrasion and ozone resistance over conventional Buna-N makes the carboxylated nitrile rubber appear promising as fuel and oil line material. In conventional Buna-N, a degree of oil resistance must be sacrificed in compounding to obtain improved ozone resistance. There may be also a slight advantage of this material over conventional Buna-N as O-ring and lip-seal material and it may extend somewhat the use of the nitrile-type polymer in automatic transmissions. Carboxylated Buna-N currently is being evaluated for such possibilities. Straight Buna-N today is borderline in a number of sealing locations. The cost of carboxylated nitrile polymer currently is $0.64 per pound. A typical compound, with a specific gravity of 1.17, costs about $0.43 per pound.

NBR-Vinyl Blends

These materials are blends of existing polymers—namely, nitrile butadiene rubber (Buna-N) and polyvinylchloride.[9,10] The Buna-N and vinyl are prefluxed during manufacture in ratios of about 70 parts Buna-N to 30 parts vinyl. The vinyl imparts to the vulcanizate principally two desirable properties deficient in Buna-N; namely, ozone resistance and flame resistance. The result is a material possessing much of the oil resistance of Buna-N, much of the ozone and weathering resistance of vinyl, and a fair degree of flame resistance. However, NBR-Vinyl compounds still support combustion unless compounded specifically for flame resistance. The flame resistance of regular Buna-N cannot be significantly improved in such manner. Additional "plus" qualities are unusually high abrasion resistance, extrudability, and colorability. On the debit side is reduced heat resistance and compression set properties at high temperatures due to the thermoplastic nature of the vinyl.

Treating some of these properties a little more quantitatively, test data show NBR-Vinyl to resist high-swell mineral oils very well. Immersion for 70 hr in ASTM Oil No. 3 at 250 F results in less than 5% swell with retention of 70% of its elongation and 90% of its tensile. Although no data are available on its properties as measured at 250 F or above, strength and tear resistance are expected to be low due to the thermo-plasticity of the vinyl in the blend. It cannot be recommended for temperatures in excess of 250 F, and in this respect loses out to straight Buna-N as a seal material in many automobile locations.

Immersion for 70 hr in ASTM Fuel B at room temperature results in about 15% swell with retention of 85% of its elongation and 65% of its tensile. This is comparable to the performance of a good, conventional Buna-N compound. The fuel resistance, coupled with ozone resistance and toughness, suggests its use as a fuel line material.

NBR-Vinyl, properly compounded, does not crack on an accelerated ozone test after two-weeks exposure to an ozone concentration of 50 pphm of air. Some compounds are reported not to crack after two weeks in an ozone concentration of 600 pphm of air. In general, the ozone resistance of NBR-Vinyl is considered comparable to that of neoprene.

It is particularly suitable for extrusions, giving very smooth, shiny surfaces.

Its high gum strength (2500 psi) eliminates the need for carbon black reinforcement to obtain respectable strengths and makes possible the production of a wide variety of bright, stable colors. Its combined weatherability and colorability indicate potential as exposed material for sealing, wire covering, and anywhere where appearance is important. A typical, high-grade, colored extrusion compound may have a tensile of over 2800 psi and elongation of over 450%, indicating that toughness also is available in this material. As a colored weatherstrip material, it is expected to be superior to vinyl in performance and at a price competitive with that of vinyl.

An excellent nonautomotive application of NBR-Vinyl is in shoe soling which takes advantage of its excellent abrasion resistance, flexing properties, weather resistance, and resistance to oils and chemicals to make it suitable for all conditions of wear. Another current use is in the cover stock of service station pump hose for similar reasons.

The price of the polymer-blend is about $0.46 per pound. Compound cost, with a specific gravity of about 1.2, may be about $0.30 per pound.

This material is available under the names of Paracril OZO and Hycar Polyblend (Hycar 1203).

Polyacrylic Rubber

The most widely used polyacrylic rubber, commonly referred to as PA rubber, is a copolymer of 95% ethylacrylate and 5% 2-chloroethylvinyl ether. The molecular structure of this polymer is shown in Fig. 12.

PA rubber is a heat- and oil-resisting material. It is significantly superior to Buna-N in its resistance to heat and hot oil environments, and is successfully used as a seal material in mechanisms such as automatic transmissions in various hot-spot locations.

PA withstands continuous exposure to dry air temperatures of 300 F with little change in its elastomeric properties (Table 1). Buna-N embrittles under this condition. The upper limit of dry air temperature for continuous exposure of PA is reported at 350 F, but at this temperature, tensile and elongation changes are significant, about 85% and 50%, respectively after one-week exposure. However, the material will withstand intermittent exposure to temperatures of 400 F or above and still retain elastomeric properties.

PA is very resistant to mineral oils, as indicated by the fact that little change occurs in high-swelling ASTM Oil No. 3 after one week at 350 F.

In ASTM Oil No. 4 (92% ASTM No. 1 and 8%

[9] "Hycar American Rubber." Manual HM-1, B. F. Goodrich Chemical Co., August, 1958.
[10] "Paracril" Bulletins, Nos. 230, 231, and 219. Naugatuck Chemical Division.

Table 2 — Comparison of Properties of High-Tensile Silicone with Conventional Silicone

	Conventional Silicone	High-Tensile Silicone
Durometer (Shore A)	50	50
Tensile, psi	950	1400
Elongation, %	300	500
Tear, lb/in.	80	180
Stiffening Temperature (F)	− 60	−166
Brittle Point, F	−100	−175
Compression Set, % (22 hr at 300 F)	24	25

Table 3 — Chemical Resistance of Fluorinated Silicone

Fluid	Immersion		Durometer Change	Volume Change, %
	Days	Temperature, F		
Isooctane	7	75	− 5	+15
Ethyl Alcohol	7	75	− 2	+ 5
ASTM No. 3 Oil	3	300	− 6	+ 4
ASTM No. 3 Oil	3	400	−14	+ 5
Skydrol 500	3	212	−26	+28
Diester Lubricant (Mil-O-7808)	3	350	−19	+10
Diester Lubricant (Mil-O-7808)	3	400	Deteriorated	
ASTM Fuel "B"	3	75	− 6	+20
ASTM Fuel "B"	3	158	− 6	+18

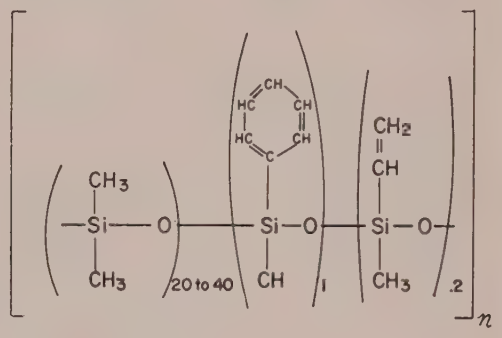


Fig. 13 — High-tensile silicone

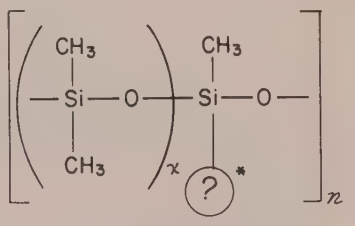


Fig. 14 — Fluorinated silicone

* Fluorinated side chain.
(Details not revealed)

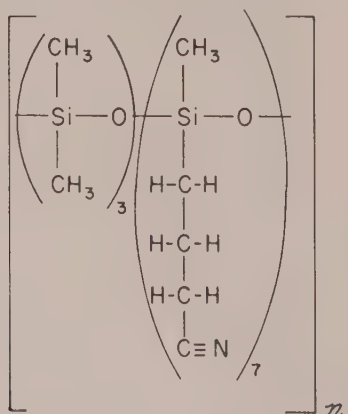


Fig. 15 — Nitrile silicone

Parapoid 10-C, chosen by ASTM to be representative of extreme pressure lubricants), PA rubber appears to be satisfactory at temperatures up to at least 300 F. Actual operating temperatures where EP lubes are used are usually considerably lower than 300 F, and at the lower temperatures PA rubber is highly satisfactory. To make another direct comparison, Buna-N embrittles in EP lubricants at elevated temperatures due to the presence of sulfur and its vulcanizing effect upon the polymer.

The resistance of PA rubber to aromatic fuels may be rated as good, but lower than that of Buna-N.

PA rubber not only resists the effects of air and most fluids at temperatures of 300–325 F, but also retains unusually high strengths at the elevated temperature. This important feature helps to explain the success of PA as a seal material in high-temperature service.

Low-temperature properties of PA are bad. The brittle point, as measured by a standard ASTM impact-type test, is usually about −5 or −10 F. Fortunately for PA rubber, the laboratory test in this case does a bad job of predicting service performance and PA continues to seal satisfactorily at temperatures down to −40 F. Apparently PA can flex sufficiently to withstand, without cracking, the rather small deformations imposed by shaft eccentricity and runout, and the heat generated rapidly elevates the temperature of the sealing lip to the point where the PA regains its true, elastomeric properties. In applications requiring considerable, sudden flexure of the component at sub-zero temperatures, as in boots and diaphragms, PA rubber cannot be used.

It is predicted that PA rubber gradually will replace more and more Buna-N as our automotive operating temperatures continue to rise. In addition to representing the next step upward from Buna-N in heat and oil resisting properties, PA also is the next step upward in price, and for that reason will be given all due consideration before adopting the more premium-priced silicones and fluorinated materials. Polymer price is $1.35 per pound. Compound cost, with a specific gravity of about 1.25, is roughly $0.70 per pound—about twice that of Buna-N. But, since in many seals the greater part of the cost is the cost of fabrication, the effect of material cost on seal cost is reduced. Consequently, finished typical PA seals usually cost only about 25% more than similar Buna-N seals.

Silicone Rubbers

Silicone rubber is well-known for its resistance to high and low extremes of temperature. Conventional silicone elastomers can be made flexible at −130 F and operative at temperatures up to 500 F. Silicone has very good compression set properties at high temperatures and low coefficient of friction against metals. It also is completely ozone resistant and has good electrical properties. These properties account for its present automotive use as spark-plug boots and rotating shaft seals in automatic transmissions. These two products, so far, are the only large-volume uses adopted by the automotive industry. Deterents to more widespread usage are high cost, low tensile and tear strengths, rather poor abrasion resistance, and deficiencies in mineral oil and fuel resistance at high temperatures. A

typical, general-purpose, silicone compound, with a specific gravity of 1.2, costs in the range of $3.25 to $3.65 per pound.

Extensive, detailed technical information on silicone rubber has been published and is readily available. This brief discussion of silicone serves merely as a background for discussions of recent developments in silicone materials—high-tensile silicone, fluorinated silicone, nitrile silicone, and RTV silicones.

High-Tensile Silicone

High-tensile silicone is a dimethyl siloxane polymer with some methyl groups replaced with phenyl and vinyl groups. The molecular structure is shown in Fig. 13.

This material was developed primarily to overcome deficiencies in strength and to retain strength at high temperatures.

Improvements in physical properties of a typical compound over a conventional, general-purpose silicone are shown in Table 2.

Resistance to mineral oils is of the same order of magnitude as regular silicone. Both swell considerably, 50% or more, when immersed in ASTM Oil No. 3 at 300 F for one week. Neither is considered resistant to fuels, particular aromatic types.

There appears to be a discrepancy between the rather poor resistance of silicone to hot mineral oils and its success as an elastomer for a front pump seal in an automatic transmission. This may be explained in part as follows: In this application silicone is used as a lip seal backed up by a garter spring. Silicone is exceptionally heat resistant so that the high temperature developed at the sealing surface does not in itself deteriorate the polymer. The mineral oil and ATF additives swell the silicone considerably, causing softening of the polymer. But this softening does not adversely affect its sealing properties. With correct lip seal design, the backing spring applies just the right amount of pressure between the silicone sealing surface and the metal surface to effect sealing. Buna-N, on the other hand, tends to harden with heat and tends to swell only slightly or even shrink due to the action of the fluid. The net result, under severe operating conditions, is embrittlement with consequent much-reduced sealing qualities.

Silicone has not been applied to O-ring or lathe-cut type seals because of its relatively low tensile strength, tear strength, and abrasion resistance. Perhaps the tougher, high-tensile silicone will prove more suitable for such applications. Its price, with a specific gravity of 1.2, is in the range of $4.50 to $5 per pound; commercially it is identified as Silastic-916.

Fluorinated Silicone

Fluorine addition to the organo-metallic elastomer affects it in much the same manner as it does the hydrocarbon elastomer—it greatly improves the resistance of the polymer to fluids, particularly the nonpolar types. Fluoro-silicone[11] is such a fluorine-containing silicone, the molecular structure of which is shown in Fig. 14. The fluorine adds fluid resistance to the polymer without materially sacrificing the other desirable properties of silicone rubber. This fluoro-silicone rubber is exceptionally resistant to swelling in mineral oils, solvents, and all types of fuels. In Table 3 are some immersion data showing its resistance to some fluids which may be of particular interest to automotive engineers. To make one direct comparison with conventional silicone, the fluorinated material swells 20% in ASTM Fuel B, whereas conventional silicone swells over 200%. It is seen from Table 3 that about the only possible automotive fluids that significantly deteriorate fluorosilicone are the highly polar hydraulic fluids and lubricants. It must be noted, however, that in these fluids the material softens and swells rather than embrittles.

Tensile strength and abrasion resistance of the fluorinated silicone are comparable to conventional silicone and still appreciably lower than most conventional organic and fluorinated hydrocarbon-type elastomers.

The exceptional oil and fuel resistance of fluorinated silicone, coupled with its extremely good high- and low-temperature properties, has resulted in early, wide usage by aircraft companies. In the future it should find use also by automotive companies when the price is reduced or when requirements for heat and oil resistance demand the use of such a premium priced polymer. Current price, depending upon volume, ranges from $18 to $24 per pound. This material is commonly known as Silastic LS-53.

RTV Silicones

RTV (room temperature vulcanizing) silicone compounds[12,13] are hydroxyl-containing silicone fluids which, with the addition of an organo-metallic catalyst, set up to a tough rubbery solid with properties somewhat comparable to conventional silicone elastomers. They are solvent-free. A number of raw compounds and catalysts are available with curing times ranging from 10 min to 24 hr.

It is difficult to visualize the use at this time of RTV silicones for large volume automotive production, but they do appear to have interesting applications, as follows:

1. Encapsulating and potting electrical items.
2. Making molds for duplication of small parts.
3. Making trial parts for prototypes and models.
4. Caulking, sealing, and patching.

The price is similar to that of conventional silicone.

Nitrile Silicone Rubber

The newest of the silicone rubbers to be introduced is NSR (nitrile silicone rubber). Structurally, NSR consists of a regular silicone polymer with a nitrile radical (—C≡N) periodically replacing a H atom of a methyl group (Fig. 15). Polarizing the silicone molecule with the nitrile groups makes the polymer more resistant to non-polar fluids such as gasoline and naphtha. The properties of the nitrile silicone are determined by the repeat frequency of the nitrile side chain and its distance from the silicon atom. The specific polymer available today, in experimental quantities only, may be identified chemically as methyl-B-cyanoethyl-dimethyl-siloxane, as shown in the figure. It might be said, in a general way, that NSR gives the oil resistance of a nitrile rubber like Buna-N and the temperature resistance of a silicone, a very desirable combination

[11] "Silastic Facts." Dow Corning Corp., May, 1957.
[12] GE Silicones Product Data.
[13] "Silastic Facts." Dow Corning Corp., December, 1958.

Table 4 — Service Life of Fluoro-Hydrocarbon at High Temperature, Time to Brittleness

Temperature, F	Life, hr
400	2400
450	1000
500	250
550	72
600	24

Table 5 — Effect on Fluoro-Hydrocarbon of Air Aging
(28 days at 450 F)

	Original	Retained, %
Tensile	2650	80
Elongation	180	70
Hardness	71	+ Points

Table 6 — Effect on Fluoro-Hydrocarbon of Oil Aging in ASTM No. 3
(1 week at 300 F)

Volume Change, %	+4
Tensile Retained, %	95
Elongation Retained, %	100
Hardness Points Change	−1

Table 7 — Relative Oil and Fuel Resistance of Elastomers

	ATF			EP-Lubes			Synthetic Hydraulic Oil[a]	Aromatic Fuel[b]
	200	300	350	200	250	300	250	R.T.
Natural or Isoprene Rubber	F	F	F	F	F	F	C	F
SBR (GRS)	F	F	F	F	F	F	B	F
Butyl (Conventional)	F	F	F	F	F	F	A	F
Chlorobutyl	F	F	F	F	F	F	A	F
High-Temperature Butyl	F	F	F	F	F	F	A	F
Chlorosulfonated Polyethylene	C	F	F	C	D	F	B	D
Neoprene	C	D	F	D	F	F	B	C
Urethane (Ether)	C	D	F	C	D	F	F	B
Urethane (Ester)	D	F	F	D	F	F	F	B
Buna-N (Conventional)	A	C	F	C	F	F	C	B
NBR-Vinyl	B	F	F	D	F	F	D	B
Carboxylated Buna-N	A	C	F	C	F	F	C	B
Polyacrylate	A	A	C	A	A	B	D	C
Silicone (Conventional)	B	BC	BC	B	C	F	AB	D
High-Tensile Silicone	B	C	C	B	C	F	AB	D
Nitrile Silicone	A	A	A				F	C
Fluoro-Silicone	A	AB	AB	A	B		F	B
Fluoro-Hydrocarbon	A	A	A	A	A	A	F	A

[a] UCON-F8B.
[b] ASTM Fuel "B."

that overcomes a major deficiency of conventional silicones.

Performance data on NSR are limited at this time. NSR currently is being evaluated by rubber companies and end users. Meanwhile, the manufacturer offers laboratory test data,[14] some of which are shown in Figs. 16 and 7, showing fuel resistance and oil resistance, respectively. Laboratory immersion data indicate marked superiority over conventional silicone in its resistance to fuels and oils. But like most of the other silicones, the nitrile silicone is somewhat deficient in tensile and tear properties.

NSR currently is being aimed principally at aircraft applications and oil seals for automatic transmissions.

Present introductory price of NSR compound with a specific gravity of 1.2 to 1.5 is $12 per pound. The projected, large-volume price is expected to be just about twice that of conventional silicone, or $6.50 to $8 per pound.

Fluoro-Elastomers

The family of fluoro-elastomers is noted for exceptional resistance to fluids at high temperatures. Before the introduction of the fluorinated materials, silicone stood alone as a usable elastomer for temperatures of 400–500 F. The silicones could withstand the temperatures but were deficient in resistance to many fluids—mineral oils and gasolines, particularly the aromatic type. The fluoro-elastomers combine good fluid resistance with good thermal stability.

Following is a brief discussion of the better known materials, including some engineering and applications information. Another fluoro-elastomer, the fluorinated silicone, was discussed in the Silicone section of this paper.

Fluorinated Hydrocarbon

This particular type of fluorinated hydrocarbon is a copolymer of vinylidene fluoride and hexafluoropropylene. The molecular configuration appears in Fig. 18.

Its excellent heat resistance is shown in Table 4, which indicates the approximate service life (time to brittleness) at various high temperatures.

Another indication of its retention of physical properties after exposure to elevated temperatures is shown by the data obtained on a typical compound in Table 5. Temperatures in the neighborhood of 300 F have no significant effect on it.

Fluoro-hydrocarbon's extreme resistance to high-swell type mineral oils is shown by physical-property changes after aging for one week in ASTM No. 3 at a realistic automotive operating temperature of 300 F (Table 6).

This elastomer is practically unaffected by ozone, oxygen, and weathering. Exposure under stress for 400 hr to the abnormal ozone concentration of 10,000 pphm causes no cracking. By way of comparison, it may be pointed out that an ozone-resistant SBR material for weatherstrips is considered satisfactory if it exhibits minute cracks after 70 hr in an ozone concentration of 50 pphm.

Low-temperature properties are moderately good. Although the material stiffens rapidly at sub-zero temperatures, brittleness is not reached above – 40 F, the critical temperature for most automotive applications.

In comparison with conventional elastomers, it exhibits very low compression set at elevated temperatures.

This copolymer has demonstrated excellent resistance to both transmission and rear axle "EP" oils, and to combination rear axle and transmission oils. Specimens immersed for 7 days at 350 F swell, in all cases, less than 5%. Tensile strength and extensibility are 80–100% retained. Hardness is virtually unaffected.

It currently is used as a valve-stem seal for a popular super-duty truck engine as a guarantee against trouble in that location.

It is practically unaffected by gasoline—even fuels

[14] "Development and Properties of Nitrile Silicone Rubber," by Dr. B. A. Bluenstein, General Electric Co.

Table 8 — Relative General Properties of Elastomers

	Tensile	Tear	Abrasion	Heat	Cold Stiffness	Cold Brittleness	Ozone
Natural or Isoprene Rubber	AB	B	B	CD	B	B	D
SBR (GRS)	B	BC	B	C	BC	B	D
Butyl (Conventional)	C	B	BC	BC	C	BC	BC
Chlorobutyl	C	B	B	BC	C	BC	AB
High-Temperature Butyl	C	B	—	B	—	—	B
Chlorosulfonated Polyethylene	BC	C	B	BC	C	B	A
Neoprene	B	B	B	C	C	BC	B
Urethane (Ether)	A	A	A	CD	BC	A	A
Urethane (Ester)	A	A	A	C	D	A	A
Buna-N (Conventional)	BC	BC	BC	B	BC	BC	D
NBR-Vinyl	BC	B	A	CD	C	C	B
Carboxylated Buna-N	AB	B	A	C	BC	B	CD
Polyacrylate	C	C	C	B	D	D	AB
Silicone (Conventional)	D	D	D	A	A	A	A
High-Tensile Silicone	C	C	CD	A	A	A	A
Nitrile Silicone	D	D	CD	A	A	A	—
Fluoro-Silicone	D	D	D	A	A	A	A
Fluoro-Hydrocarbon	BC	BC	B	A	D	BC	A

that are better than 50% aromatics. Rubber parts for fuel and gasoline pumps are a natural for this material. Carburetor cups and seals currently are being made from it.

This fluorinated material has outstanding resistance to silicate esters, is less resistant to phosphate esters, and is not considered satisfactory for highly polar fluids or for organic acids, ketones, aldehydes, and the like.

It is obvious that such fluorinated materials have the wherewithal to handle tough automotive diaphragm and sealing problems as future requirements become increasingly severe. However, there is one very good reason why the industry does not solve most of its problems, current and future, of seals, gaskets, tubing, and coated fabrics with the stroke of a pen specifying fluoro-elastomers. That reason is price. The polymer sells for about $15 per pound. Compounded with magnesium oxide and carbon black, the price of a typical compound, with a specific gravity of about 1.9, is about $10.50 per pound. Typical of these materials are Viton and Fluorel.

Elastomer Comparisons

Direct comparisons of the foregoing elastomers are made in Tables 7–9.

Oil and Fuel Resistance

Table 7 shows in a general way the type of performance that can be expected at realistic operating temperatures of typical, high-quality compounds of each polymer in automatic transmission fluids, rear axle "EP lubes," one synthetic hydraulic oil (UCON-F8B) of some current interest to automotive companies, and a fuel of high aromatic content, ASTM Fuel B. The relative ratings are defined as follows: A — Exceptional, B — Good, C — Fair, D — Poor (not recommended), F — Completely unserviceable.

A double letter is used at times to permit a somewhat finer comparison and indicates a position more or less between the two major ratings.

General Properties

Table 8 similarly shows a comparison of these

Fig. 16 — NSR fuel resistance

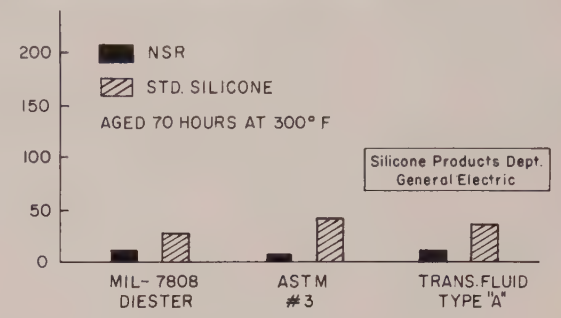


Fig. 17 — NSR oil resistance

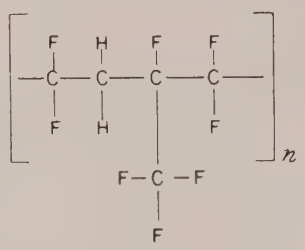


Fig. 18 — Fluoro-hydrocarbon (vinylidene fluoride-hexafluoropropylene)

elastomers with respect to several additional properties usually of great importance in automotive applications of these elastomers. For these properties, the F rating is omitted. Consequently, the worst material for a given property is rated "poor" but not "completely unserviceable," since this depends totally upon the severity of service use. It must be emphasized that these ratings are general, that they apply to more or less typical high-grade compounds. It is fully recognized that a property rating of a material may be altered through compounding, that by special or "trick" compounding certain properties can be accentuated, but usually at the sacrifice of some other desirable property. The ratings assigned to the various materials apply to practical compounds with good balance between important physical properties. In fact, the ratings assigned many of the elastomers are lower than those usually associated with the elastomer. This is due to the fact that the ratings are relative and generally have been shifted downward due to the competition offered by many of the "super" materials available today. For example, the abrasion resistance and low-temperature properties of natural rubber in the past have been considered "excellent" but today they are only "good," since natural rubber must take a backseat to urethane in abrasion resistance and to the silicones in low-temperature flexibility. Similarly, the fluorinated materials have far surpassed the more conventional

elastomers like neoprene and Buna-N in resistance to mineral oils.

Cost

Most interesting to automotive engineers, always forced to face the hard facts of economics, is the cost picture presented in Table 9. Published cost per pound can be misleading due to the wide variance in specific gravity of the materials. A more realistic comparison is made on the basis of cost per pound-volume, the product of cost per pound and specific gravity, which gives the cost comparison on an equal volume basis. Even polymer cost per pound-volume can be somewhat misleading to the application or design engineer because of the fact that finished compound cost usually differs from basic polymer cost due to the incorporation of compounding ingredients. The last column, therefore, is included to present in a realistic manner a comparison of the prices he might expect to pay for typical high-grade compounds, ready for use, on an equal volume basis.[15] Elastomers currently specified for use in the automotive industry range in cost from the twenty-cent material to the twenty-dollar material. In fact, some weatherstrip materials, less expensively compounded than those used in the comparisons, are obtainable for less than $0.15 per pound-volume. Fluoro-hydrocarbon, currently used as a valve-stem seal, probably represents the most costly elastomer currently used at $20 per pound-volume.

Table 10 lists the trade names of the elastomers discussed in this paper.

Table 9 — Costs of Polymers and Compounds[a]

	Polymer cost/lb	Polymer Specific Gravity	Polymer cost/ lb-vol	Typical Compound[b] cost/ lb-vol
Natural or Isoprene Rubber	$ 0.30	0.92	$ 0.28	$ 0.25
SBR (GRS)	0.24	0.93	0.22	0.21
Butyl (Conventional)	0.24	0.92	0.22	0.21
Chlorobutyl	0.30	0.92	0.28	0.26
High-Temperature Butyl	0.24	0.92	0.22	0.23
Chlorosulfonated Polyethylene	0.70	1.10	0.77	0 63
Neoprene	0.39	1.23	0.48	0.33
Urethane	1.15	1.06	1.22	1.30
Buna-N (Conventional)	0.50	0.98	0.49	0.40
NBR-Vinyl	0.46	1.06	0.49	0.35
Carboxylated Buna-N	0.64	1.00	0.64	0.50
Polyacrylate	1.35	1.09	1.47	.90
Silicone (Conventional)	3.25[c]	1.20	3.90	3.90
High-Tensile Silicone	4.75[c]	1.20	5.70	5.70
Nitrile Silicone	12.00[c]	1.35	16.20	16 00
Fluoro-Silicone	19.00[c]	1.42	27.00	27.00
Fluoro-Hydrocarbon	15.00	1.82	27.30	20.00

[a] Costs are approximate. For exact costs consult supplier or latest price list.
[b] Comparable high quality level.
[c] Cost of compound rather than polymer, as supplied by manufacturer.

Table 10 — Trade Names versus Chemical Names of Elastomers

Isoprene Rubber	Ameripol SN
	Coral Rubber
	Natsyn Rubber
	Shell Polyisoprene
Chlorobutyl	MD-551
Chlorosulfonated Polyethylene	Hypalon
Urethane	Adiprene-C
	Adiprene-L
	Chemigum-SL
	Daykollan
	Estane
	Genthane
	Vibrathane
	Vulcollan
NBR-Vinyl	Paracril OZO
	Hycar 1203 (Hycar Polyblend)
Carboxylated Buna-N	Hycar 1072
Polyacrylate	Hycar 4021
High-Tensile Silicone	Silastic-916
Fluoro-Silicone	Silastic-LS-53
Fluoro-Hydrocarbon	Viton
	Fluorel

Conclusion

It is hoped that this paper makes clear the fact that elastomers are available today to solve most of the current technical problems of rubber in the automotive industry. The cost of many, however, seems prohibitive. The result is that they are used only as a last resort—usually to solve an otherwise insoluble or extremely serious problem. Volume use of the premium-priced polymers will reduce the price, but such materials as the fluorinated elastomers, because of the basic materials and processes involved, always will be expensive in comparison with the more conventional, widely used, elastomers. But certainly there is a place for premium-priced elastomers in the automotive industry. Seals of polyacrylates, silicones, or fluorinated hydrocarbons may permit higher temperature operation of automatic transmissions and eventually may eliminate the expensive oil cooler on many transmissions. Mechanical rubber goods may take advantage of the extreme toughness, strength, and abrasion resistance of urethane rubber and, by letting a small volume of rubber do the work of a relatively large volume of conventional elastomer, possibly actually reduce cost and save space as a bonus feature. The automotive industry has not yet taken full advantage of all the superior and peculiar properties of the new elastomers. To accomplish this, components must be designed around the material in contrast to the more normal procedure of fitting a material to an established design. A sound approach in many instances would be to break with precedent and obsolete the design, apply imagination, and use the elastomer properly. The results may be astounding.

DISCUSSION

Comments on Polybutadiene and Butyl Acrylate-acrylonitrile

— Gust A. Ilkka
General Motors Corp.

MR. SCHMUCKAL indicated in his talk that polymer manufacturers have been reluctant to publish very much about the deficiencies of their materials. While this statement is essentially correct, I think this situation has improved in recent years, as compared to many years ago when practically no data on polymer shortcomings were published. Actually, it would seem to be an advantage to the polymer makers to reveal known weaknesses of their materials in order to prevent misapplications and the resulting adverse publicity.

As a supplement to Mr. Schmuckal's paper two addi-

[15] A pound-volume represents the volume occupied by 1 lb of a material with a specific gravity of 1.0 — roughly 1 pt.

tional elastomers should perhaps be mentioned here. These are: polybutadiene and the butyl acrylate-acrylonitrile copolymer. Polybutadiene can be made in several cis-trans configurations with varying properties. The polybutadienes can be made with existing plant facilities and are potentially lower cost materials than the polyisoprenes. fifty-fifty blends of cis 1,4-polybutadiene with natural rubber are claimed to be equivalent to or better than natural rubber alone for tire use. In this manner, polybutadiene could help alleviate shortages of natural rubber and stabilize its price.

The second elastomer is a copolymer of butyl acrylate and acrylonitrile. It was formerly known as Acrylon BA-12 and was used to some extent in seal applications until its manufacture was discontinued. This BA-12 material is now being produced by another company at a somewhat higher price of $1.75 per pound. This BA-12 polyacrylate is quite similar in properties to the Hycar 4021 polyacrylate discussed by Mr. Schmuckal except that it has about 10–20 F better low-temperature properties than the Hycar 4021.

Of all the elastomers discussed by Mr. Schmuckal in his paper, the polyurethanes seem to be the most intriguing in that so many different types are available with such a wide range of mechanical properties. Some can be cast as liquids, others can be molded like conventional rubber compounds, while still others are thermoplastic and should be adaptable to injection molding.

It is expected that the discovery of new types of elastomers will become accelerated now that the polymer scientist has various block, graft, interfacial, and stereospecific polymerization techniques at his disposal and more and more new monomer materials are being developed. So, the automotive engineer can look forward to a still more hectic but interesting life as far as elastomers are concerned.

Automotive Use of Elastomers Depends upon Improved Design

— W. J. McCortney
General Tire & Rubber Co.

VOLUME-WISE, as the cost increases over a dollar a pound, the potential use in the automotive industry of special purpose rubbers decreases to a point that such polymer development is unprofitable unless other supporting uses in other industries can be found. Nevertheless, you will note from Mr. Schmuckal's paper that uses in small amounts of these high-priced rubbers can be extremely important to the functioning of a modern automobile. I am certain that automotive-wise there are many excellent polymers that will remain on the shelf until broad enough usage can be found for them. On the other hand, there are many low-cost elastomers in various stages of development that the automobile will use with or without supporting volume from other industries. Examples of these are the cis-polybutadiene, the amorphous olefinic, and various graft polymers. All are potentially of low cost and could readily be made available in almost unlimited quantities.

The use of new elastomers in very many instances is dependent on further design and improved accompanying materials. Units such as automatic transmissions and hydraulic brakes could be made either smaller or more efficient if increased temperatures were possible. While elastomers for such units are now available, compatible fluids to withstand higher temperatures are in the doubtful class.

The 100,000-mile tire has been imminent for several years but each year the service becomes more severe and we have been finding it difficult to maintain high abrasion resistance, high-speed performance, and low noise level in the smaller rim diameter tires now in use. It is my opinion that design will contribute much toward improved tire performance and may make it possible to use elastomers not suitable for conventional tires. This has been most recently proved by redesign of the tire for butyl rubber.

I have mentioned both the cost of new elastomers and the 100,000-mile tire. Such a tire would undoubtedly cost much more than the present ones due to the necessity of using premium elastomers. How much more will the car buyer pay for such tires?

Mr. Schmuckal mentions that urethane rubbers fall into the harder Shore A durometer ranges. I would like to modify this by saying that by using Genthane-S with its different cross linking mechanism, urethanes can be compounded to 40 and 50 durometer.

Lower Quality Elastomers Suitable For Automotive Applications

— R. W. Malcolmson
E. I. du Pont de Nemours & Co.

THIS PAPER is unquestionably one of the most comprehensive reviews of elastomers ever to be presented to automotive engineers. Copies of it for reference are almost a must on the desks of materials engineers; ready availability to all automotive design engineers would be desirable.

Discussion of two older, well-known elastomers, neoprene and SBR (formerly known as GR-S, a general-purpose, nonoil resistant synthetic rubber), is all that is lacking to make the coverage really complete. Of course, these elastomers are included in the tabular comparisons of oil and fuel resistance (Table 7), general physical properties (Table 8), and cost (Table 9).

Now, a word about these three tables. I'd like to reemphasize the pitfalls of qualitatively rating such a wide variety of materials with relatively few grade letters. For example, letter ratings from A through D cannot adequately describe the vast differences in oil and fuel resistance between the fluoroelastomers, such as Viton, and chlorosulfonated polyethylene (du Pont's Hypalon). Similarly, the ozone resistance ratings in Table 8, to take another example, do not adequately describe the difference between Hypalon, which is completely ozone-proof at all concentrations so far as we can determine, and some of the less resistant elastomers such as NBR and SBR. Nevertheless, these ratings can be useful with proper appreciation of the drawbacks noted above. It's another instance of where "A little knowledge is a dangerous thing."

The subject of relative cost is perhaps even more dangerous in the hands of an untrained reader of this paper. Mr. Schmuckal has been careful to point out that the pound-volume costs are for comparable high-quality compounds. However, the bulk of automotive applications do not require the high level of quality represented by the prices quoted. In the case of neoprene, for example, pound-volume costs for the most widely used compounds are in the $0.18–0.25 range (25–45% lower than indicated in the table). In the case of other materials (many of the silicone elastomers and the castable urethanes) little or no latitude is available to rubber goods manufacturers to adjust the compounds to cost or quality levels other than those for which materials costs are shown. Hence, the user of this table must realize the differences among various elastomers in this respect.

Regarding some of the specifics of Mr. Schmuckal's paper, I'd like to echo one of his remarks on urethane rubber: "Ths material, being so different from conventional elastomers, should not be considered a direct replacement for currently used materials. To take full advantage of its superior properties, the component must be designed around the materials. . . . The resultant savings in rubber volume in many cases would offset the higher cost per pound and give additional bonus features in reduced weight and space requirements." In addition, the spectrum of urethane polymers is so wide that they must be considered not only new elastomers, but must be regarded as new design materials in the harder range where they will compete with such materials as nylon, phenolics, epoxies.

Engine Rumble

—a barrier to higher compression ratios?

R. F. Stebar, W. M. Wiese, and R. L. Everett
General Motors Research Laboratories

A NEW VARIETY of engine noise has become apparent during the past few years — engine rumble. Rumble is a low-pitched, rapping noise which is associated with combustion in high compression engines. The noise is very loud and harsh in severe cases and might be compared to the noise from loose connecting-rod bearings. Rumble is most likely to occur at high engine speeds and loads such as might be encountered in passing or hill climbing.

At the present time, rumble is recognized as a barrier that must be overcome before substantially higher compression ratios and correspondingly higher engine efficiencies can be reached.[1,2] Rumble is objectionable not only from a noise standpoint but also because the combustion process is abnormal and uncontrolled. The abnormal combustion associated with rumble can result in reduced engine efficiency.

This paper describes the phenomena of engine rumble both from the noise and the combustion standpoints. More important, however, studes were made to determine the contributions of oils, fuels, and engine operating variables to the occurrence and control of rumble.

Rumble Phenomena

Characteristics of Rumble Sound

The sound of engine rumble is extremely objectionable to passengers inside the car. Rumble can be distinguished from combustion knock by its lower pitched sound and by the fact that it usually occurs at high engine speeds and loads. An investigaiton of the characteristics of the sound revealed that rumble was the result of resonant engine vibrations in the frequency range 500–1600 cps.

Sound recordings were taken inside the engine compartment of a 12/1 compression ratio car under both normal and rumbling conditions. These recordings were made under full-throttle conditions at an engine speed of 3000 rpm. Frequency analyses of these recordings are shown in Fig. 1. Relative sound intensity is shown as a function of frequency in cycles per second (cps). The upper curve represents the sound analysis for this engine during rumble; the lower curve represents the sound analysis during normal operation. The two curves differ principally in two respects: (1) the overall sound intensity for the rumbling engine at a higher level than for the normal engine from about 500 to 1600 cps, and (2) in particular, the sound intensity is considerably higher at about 800 and 1000 cps for the rumbling engine. These frequencies vary slightly for different engines. As a matter of interest, the frequency of knock for this engine was about 5500 cps.

It was also observed from sound analyses that the characteristic rumble frequencies (800 and 1000 cps) did not change when engine speed was varied. This observation indicates that the rumble sound is the result of resonant vibrations occurring in the engine structure.

The foregoing conclusion was confirmed by an investigation of main bearing cap deflection. For these investigations strain gages were mounted on the bottom surface of the main bearing caps. The deflections of the bearing caps were recorded during both normal and rumbling operation. Typical oscillograph recordings of bearing cap deflections are shown in Fig. 2. It is evident that vibrations exist in the crankshaft main bearing area during

[1] "More Efficient Utilization of Fuels," by C. F. Kettering. SAE *Quarterly* Transactions, Vol. 1, October, 1947, pp. 669–680.
[2] "New Look at High Compression Engines," by D. F. Caris and E. E. Nelson. SAE Transactions, Vol. 67, 1959, pp. 112–124.

A NEW BARRIER to higher compression ratios has recently become apparent — engine rumble! This phenomenon will prevent further increases in compression ratio unless corrective measures are taken.

This paper describes the phenomena of engine rumble not only in terms of the noise and vibrations that emanate from the engine but also in terms of the pressure development inside the cylinder. Rumble is the result of abnormally rapid pressure buildup in the combustion chamber due to multiple ignition of the fuel-air mixture by glowing deposits.

Since deposits are responsible for the occurrence of engine rumble, studies have been made to determine the contribution of various gasolines and oils to the rumble tendency of the deposits formed. Results from dynamometer and road tests show that combustion-chamber deposits formed by the use of some oils and fuels are considerably less likely to cause rumble than deposits from others. A phosphorus gasoline additive was tested at concentrations from ½ to 1 theory to determine its effectiveness in reducing rumble in 12/1 compression ratio cars.

Another aspect of the rumble problem is the inherent resistance of the fuel itself to ignition by deposits. Deposit ignition resistance ratings of several fuels were made in a car on the road.

Several engine design and operating variables were studied to determine their effect on the occurrence of rumble. These included compression ratio, air/fuel ratio, inlet air temperature and humidity, coolant temperature, engine load, and engine speed.

Evidence is shown that engines having compression ratios as high as 12/1 can be operated satisfactorily with respect to rumble if fuels and oils are carefully selected.*

rumble which do not exist during normal operation. This is in agreement with Starkman and Sytz[3] who have concluded that crankshaft vibration is the main source of the rumble noise.

Cause of Rumble

Rumble is caused by abnormal combustion which is directly related to the presence of combustion chamber deposits. This can be demonstrated by the observation that rumble is eliminated by deposit removal. Furthermore, rumble can be initiated by introducing deposits into the carburetor airstream.

Detailed studies of the combustion process during rumble were made by instrumenting an engine so that pressure development in one cylinder could be observed. A vibration pickup was mounted on the crankcase near the instrumented cylinder to detect the occurrence of rumble. Shown in Fig. 3 are reproductions of actual pressure-time diagrams recorded during normal combustion and during rumble. For comparison, a pressure-time diagram obtained under knocking conditions is also shown. The engine condition for all three diagrams was 3000 rpm — full throttle.

A comparison of the rumble and normal diagrams shows that cylinder pressure departs from normality very abruptly during the compression stroke. At this point hot combustion-chamber deposits have initiated several flame fronts in addition to the one started by the spark discharge. The result is an abnormally rapid release of energy. These events would be analogous to igniting the charge with several spark plugs located at various positions in the chamber. The abnormal characteristics of the rumble pressure-time diagram may be described as

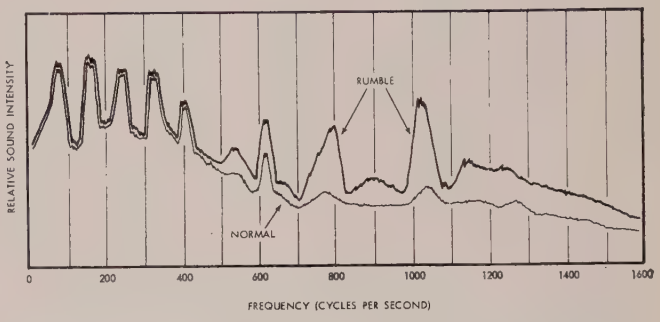


Fig. 1 — Frequency analyses of engine noise during normal and rumbling operation (12/1 compression ratio, 3000 rpm, full throttle)

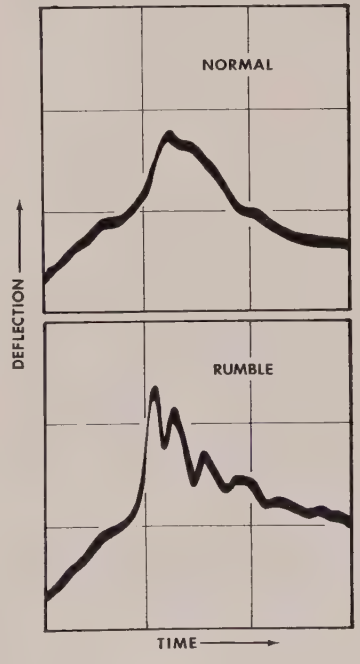


Fig. 2 — Main bearing cap deflection during normal operation and during rumble (12/1 compression ratio, 2500 rpm, full throttle; deflection of No. 4 main bearing cap during firing of No. 5 cylinder)

* Paper presented at SAE National West Coast Meeting, Vancouver, B. C., August 10, 1959.

[3] "Identification and Characterization of Rumble and Thud," by E. S. Starkman and W. E. Sytz. SAE Transactions, Vol. 68, 1960, pp. 93–100.

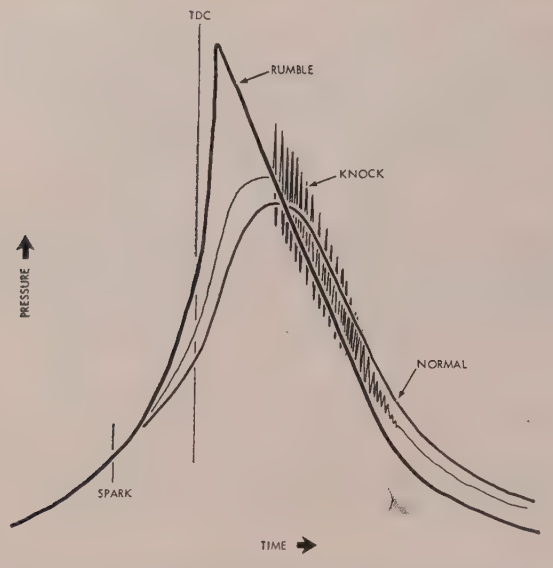


Fig. 3 — Pressure-time diagrams for normal, rumbling, and knocking combustion (12/1 compression ratio, 3000 rpm, full throttle)

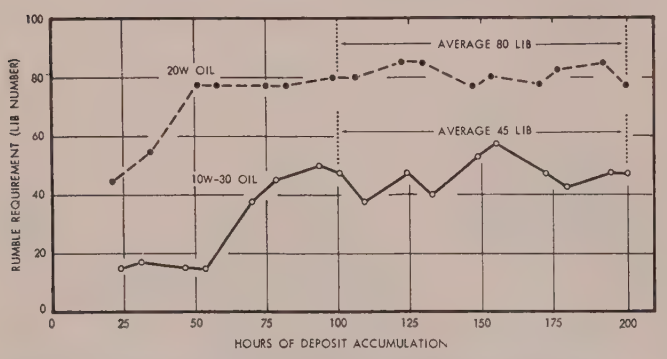


Fig. 4 — Engine rumble requirements during deposit accumulation (10/1 compression ratio, 3000 rpm, full throttle, engine dynamometer)

follows: (1) a very rapid rate of pressure rise, (2) a very high peak pressure, (3) a sharp peak, and (4) the occurrence of peak pressure earlier in the cycle. Engine cycles having these characteristics cause shock excitation of the engine parts which results in rumble.

It should be emphasized that noise is not the only objectionable characteristic of rumble. Engine efficiency is reduced as a result of the abnormal combustion which accompanies rumble. Calculations of indicated work made from the pressure-time diagrams in Fig. 3 show that 5–6% less work was done during the rumble cycle than during the normal cycle. Results from other pressure-time diagrams indicate losses as great as 13%. The lost work is in the form of heat which is rejected to the combustion-chamber walls. This rejected heat may result in overheated and damaged engine parts, although no engine failures have been encountered that can be directly attributed to rumble.

Rumble Mechanism

The mechanism of rumble discussed in the preceding paragraphs may be summarized as follows:

1. Multiple ignition of the fuel-air mixture by engine deposits.
2. Abnormal pressure development.
3. Shock excitation of engine parts.
4. Resonant vibration producing rumble noise.

Table 1 — Fuel Analysis Data

	Fuel		
	A	B	C
Specific Gravity at 60 F	0.732	0.750	0.743
Tetraethyllead, ml/gal	3.09	2.74	nil
Reid Vapor Pressure, lb	4.8	12.0	12.3
Phosphorus, theories	nil	0.22	nil
Motor Octane No.	100	88	87
Research Octane No.	110	100	99
Distillation, F			
initial	117	84	85
10%	171	110	118
50%	214	207	214
90%	248	318	324
end	317	396	379
Hydrocarbon Content			
Saturates, %	72	53	52
Olefins, %	10	16	16
Aromatics, %	18	31	32

Rumble Measurement Technique

The occurrence of rumble was measured through the use of leaded isooctane-benzene (LIB) reference fuels.[4] The LIB reference fuel service is made up of blends of leaded isooctane in leaded benzene. Each blend contains 3 ml tel/gal. The LIB number of the individual blends corresponds to the percentage of leaded isooctane in the blend. Since isooctane has very high resistance to deposit ignition and benzene has very low resistance,[5] blends of the two fuels provide a series of reference fuels having varying deposit ignition resistance. Reference fuels were blended in 10% increments.

The LIB fuels are used to measure the deposit ignition or rumble requirement of an engine in much the same manner as the octane scale is used to measure antiknock requirement. The rumble requirement of an engine is defined as the LIB number of the fuel which produces trace rumble. Knock is not encountered with the LIB fuels because all blends have Research and Motor octane ratings of 115 or higher.

Crankcase Oils — Effect on Rumble

During normal engine operation, crankcase oil is drawn past the piston rings and through the valve guides into the combustion chamber of an engine. Although the amount of oil consumed in this manner is small, the effect on deposit ignition and, hence, rumble occurrence is extremely important. Certain oils form combustion-chamber deposits which cause more rumble than other oils.

The relative contribution of commercial oils to the occurrence of rumble was explored in two 1959 production V-8 engines at 10/1 compression ratio using a dynamometer test technique. Equilibrium combustion-chamber deposits were accumulated for each oil on a light-duty deposit buildup schedule. A commercial gasoline (Fuel B, Table 1) which contained 0.22 theories[6] of phosphorus was used

[4] These reference fuels were originally proposed by the Coordinating Research Council for use in deposit ignition studies.
[5] Private report, Motor Combustion Chamber Deposits Group, Coordinating Research Council, Inc.
[6] One theory of phosphorus is the theoretical amount required to convert all of the lead in the fuel to lead orthophosphate.

throughout the oil evaluation tests.

Engine rumble requirements were determined periodically at 3000 rpm — full throttle. Details of both the rating procedure and deposit accumulation schedule can be found in Appendix I. To illustrate how oil rumble ratings were obtained from the test results, data for two oils are shown in Fig. 4. Engine rumble requirements are shown as a function of deposit accumulation time. The individual requirements between 100 and 200 test hr were averaged to provide a stabilized rumble rating for each oil; the higher this rating, the more undesirable the oil with respect to rumble. For example, use of the 20W oil rated at 80 LIB is more likely to cause engine rumble than use of the 10W-30 oil rated at 45 LIB numbers. Reproducibility of test results was considered statisfactory since rumble ratings of these two oils repeated within 5 LIB numbers in duplicate tests.

The important influence of crankcase oils on engine rumble is shown in Fig. 5. Included are data for ten commercial MS oils of 20W, 10W, and 10W-30 SAE viscosity classifications. The rumble ratings of these oils differed widely. For the five 20W oils, rumble ratings ranged from 65 LIB for Brand D to greater than 100 LIB numbers for Brand A. Ratings for the four 10W-30 oils ranged from 15 LIB for Brand A to 55 LIB for Brand E. The one 10W oil tested, Brand D, was rated at 55 LIB numbers. *In ever case the 10W-30 oils had lower rumble ratings than did the 20W oils.*

Results from laboratory analyses of the ten oils (Table 2) indicated a correlation between rumble ratings and only volatility characteristics. The best correlation was found when volatility was expressed in terms of per cent residue from a 550 F circulating-air pan evaporation test (Appendix II). This correlation is shown in Fig. 6. In general, the higher the per cent residue, the higher is the rumble rating of the oil or, in other words, oils having a high percentage of residue are more likely to cause engine rumble. The pan evaporation test or a similar test may offer a means of detecting oils which will produce excessively high engine rumble requirements.

A road test evaluation of two oils, Brand C 20W

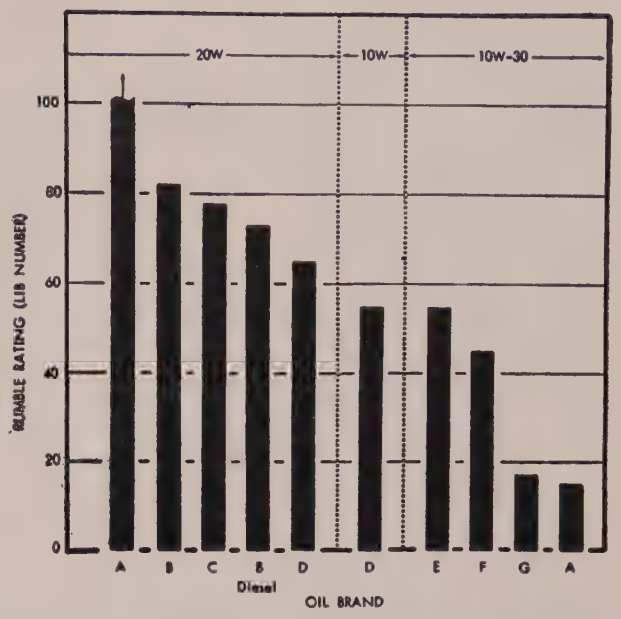


Fig. 5 — Rumble ratings of commercial MS engine oils (10/1 compression ratio, 3000 rpm, full throttle, engine dynamometer)

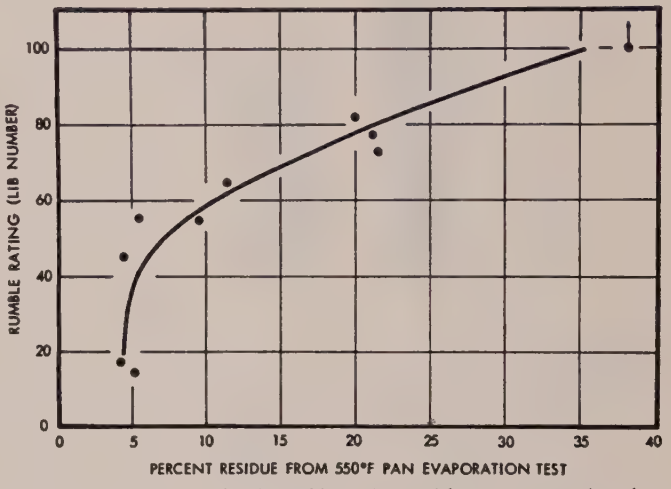


Fig. 6 — Correlation of oil rumble ratings with pan evaporation data

Table 2 — Analytical Data for Commercial MS Oils

	Brand A 20W-30	Brand B 20W	Brand B 20W (Diesel)	Brand C 20W	Brand D 20W	Brand D 10W	Brand E 10W-30	Brand F 10W-30	Brand G 10W-30	Brand A 10W-30
Pan Evaporation Data, % Residue at 550 F	38.1	20.0	21.0	21.1	11.5	5.4	9.4	4.4	4.1	5.2
Distillation, F at 10 mm Hg vacuum										
10%	448	480	445	504	492	454	454	450	460	445
20%	471	505	470	533	515	472	472	465	465	454
30%	514	522	490	551	532	487	487	477	472	467
40%	550	537	505	567	547	500	502	485	481	480
50%	591	551	552	582	560	513	518	494	491	495
60%	635	567	542	597	573	528	536	503	501	513
70%	—	589	570	614	587	546	556	514	517	532
80%	—	627	611	637	603	567	583	528	525	552
90%	—	—	—	674	631	598	—	553	552	—
End Point	660	653	641	700	647	601	614	562	562	565
% Distilled	66.3	84.7	83.8	94.4	94.4	91.0	88.7	92.2	94.4	86.3
Viscosity, cs										
at 100 F	96.8	77.0	78.9	69.9	56.0	45.4	66.1	66.7	75.9	63.5
at 210 F	11.3	8.92	9.37	9.04	9.48	7.19	11.2	11.67	12.70	10.22
Viscosity Index	112	97	104	112	140	125	142	145	142	137
Specific Gravity at 60 F	0.876	0.885	0.899	0.876	0.882	0.876	0.878	0.862	0.873	0.872

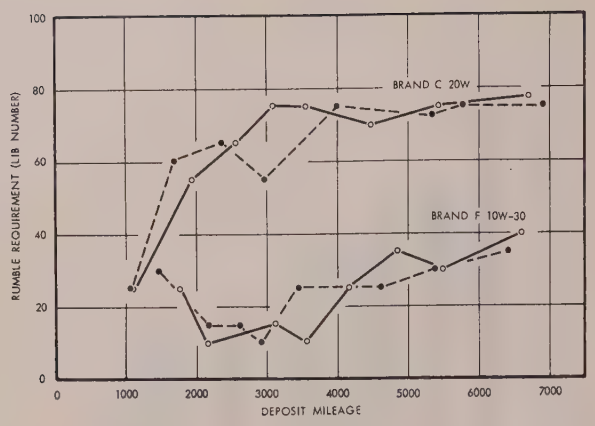


Fig. 7 — Road tests to verify dynamometer ratings of oils (10/1 compression ratio, 3000 rpm, full throttle)

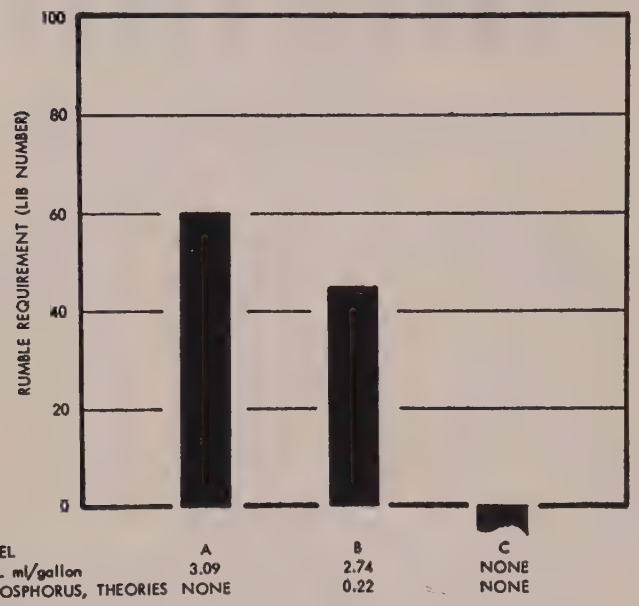


Fig. 8 — Variation in rumble requirement when deposits are accumulated with different types of fuel (10/1 compression ratio, 3000 rpm, full throttle, engine dynamometer)

and Brand F 10W-30, verified the rumble ratings of these oils determined by the laboratory tests. For the road tests, four cars having engines of the same make and model as used in the dynamometer tests were operated in company transportation service. The factory-fill oil was drained at approximately 1000 miles, and thereafter two cars were operated on Brand C 20W and two on Brand F 10W-30 oil. Engine rumble requirements were measured at approximately 500-mile intervals. The requirements were determined at 3000 rpm, full throttle. Engine speed was held constant with the brakes.

Results of the road tests are presented in Fig. 7. Engine rumble requirements are shown as a function of car mileage. After 4500 miles the cars operated on Brand C 20W oil had rumble requirements of approximately 75 LIB numbers while the cars using Brand F 10W-30 oil had requirements of approximately 35 LIB numbers. A comparison of these road test results and the laboratory results indicates that both tests rank the oils in the same order; furthermore, the rumble requirements of the cars agreed reasonably well with those obtained for the laboratory engines.

Fuels — Effect on Rumble

The fuel used in an engine can affect the occurrence of rumble in two ways: first, through the deposits formed in the combustion chamber when the fuel burns; and second, through the inherent *resistance* of the fuel to deposit ignition when exposed to the high temperatures and pressures in the cylinder during the compression stroke. Both of these effects must be considered when evaluating the merits of a gasoline.

Fuel Contribution to Deposits

Fuels as well as oils affect the type of deposits formed in the combustion chamber. The deposits formed when using certain types of fuel are more likely than others to cause deposit ignition which can result in rumble. It has been shown by past investigations that fuels containing very high-boiling hydrocarbons or certain heavy aromatic hydrocarbons form deposits which are more prone to cause deposit ignition and rumble than other fuels.[7,8,9,10] These same investigations and others[11,12] have also shown that gasoline additive treatment has a marked effect on the deposit ignition tendency of the deposits formed.

Fig. 8 illustrates the variation in engine rumble requirement that can result from the use of different types of fuel. These data were obtained on an engine dynamometer using the same 1959 production engine and test procedure as was used for the engine oil tests. Three different fuels were tested with a commercial 10W-30 oil. Deposits were accumulated with each fuel and average rumble requirements were determined. Fuel A is a gasoline containing 3 ml tel/gal. Fuel B is a commercial gasoline containing approximately 3 ml tel/gal plus 0.22 theories of a phosphorus ignition control compound. Fuel C is a commercial gasoline containing neither tel nor phosphorus. Analytical data for the three test fuels are shown in Table 1. It is obvious from Fig. 8 that tel is a large contributor to the occurrence of rumble since the engine would not rumble on the lowest LIB fuel when deposits were accumulated with Fuel C. Fuel B containing a small amount of a phosphorus ignition control compound reduced the rumble requirement of the engine considerably compared to Fuel A which contained only tel. These data serve to illustrate the effects of tel and phosphorus even though the hydrocarbon composition of Fuel A is somewhat different from that of Fuels B and C.

Phosphorus Fuel Additive

From the preceding results it would appear desirable to eliminate tel from gasoline—if rumble were the only concern. From an antiknock standpoint, of course, tel is very effective and economical; therefore, its use will undoubtedly continue. How-

[7] "Rumble — Deposit Effect at High Compression Ratios," by A. E. Felt, J. A. Warren, and C. A. Hall. SAE Transactions, Vol. 67, 1959, pp. 138–144.
[8] "If You Squeeze Them, Must They Scream?," by W. M. Wiese. SAE Transactions, Vol. 67, 1959, pp. 175–183.
[9] "Knock, Rumble, and Ping," by H. F. Hostetler and W. R. Tuuri. SAE Transactions, Vol. 67, 1959, pp. 152–157.
[10] "Knock, Knock: Spark Knock, Wild Ping, or Rumble?" by R. H. Perry, Jr. and H. V. Lowther. SAE Transactions, Vol. 67, 1959, pp. 145–151.
[11] "Improve Fuel with Phosphorus Additives," by R. E. Jeffrey, L. W. Griffith, and E. Dunning, and B. S. Baldwin. *Petroleum Refiner*, Vol. 33, August, 1954.
[12] "Some Effects of Fuels and Lubricants on Autoignition in Cars on the Road," by R. K. Williams and J. R. Landis. SAE Transactions, Vol. 62, 1954, pp. 57–71.

ever, as indicated in Fig. 8, phosphorus gasoline additives can be used to reduce rumble when leaded gasolines are used. The phosphorus combines with lead during combustion, forming lead-phosphorus compounds which are less likely to cause deposit ignition than the basic lead chlorides and bromides[13, 14] which are normally formed when leaded gasoline is burned in an engine. Because of this, use of phosphorus additives in gasoline appears to be a practical compromise; the antiknock advantages of tel can be retained while at the same time deposit ignition and, hence, rumble can be minimized.

Nearly all of the phosphorus-containing gasolines marketed today contain between 0.1 and 0.3 theories of phosphorus. This amount of phosphorus appears to be reasonably effective at today's compression ratios but may not be adequate when higher compression ratios are reached. Because of this, a deposit accumulation road test program was carried out to investigate the effectiveness of various concentrations of a phosphorus gasoline additive in reducing the rumble tendency of high compression cars.

The additive used during these tests was a commercial dimethylxylyl-phosphate compound. Tests were run at three different phosphorus concentrations; namely, ½ theory, ¾ theory, and 1 theory. The base fuel used during tests with each phosphorus concentration was a blend containing 75% alkylate, 15% toluene, and 10% diisobutylene, plus 3 ml tel/gal. This fuel is the same as Fuel A shown in Table 1.

Four 1958 test cars equipped with special 12/1 compression ratio V-8 engines were used during these tests. Cars of two makes were involved. Compression ratios of the engines were increased through the use of special pistons and/or cylinder heads obtained from the respective manufacturers. Combustion-chamber deposits were accumulated under suburban-highway driving conditions with vehicle speed limited to 55 mph. LIB rumble requirements were determined at 200-mile intervals. These requirements were obtained during full-throttle accelerations from approximately 30 to 70 mph. A mid-continent fully distilled 20W oil was used.

The tests were arranged so that all concentrations of phosphorus were tested simultaneously in different cars during a given time period. After stabilized rumble requirements had been obtained, the test fuels containing various concentrations of phosphorus were switched to different cars. The tests were repeated until each concentration of phosphorus had been tested in each car. Deposits were accumulated for 3000–5000 miles with each fuel — until a stabilized rumble requirement was obtained. Deposits were not removed from the combustion chambers between tests. Tests were also run to determine the rumble requirement of each car when no phosphorus was used in the tank fuel.

Results of these tests are shown in Fig. 9. LIB rumble requirement is shown as a function of the concentration of phosphorus additive used in the tank fuel. Each point represents the average of requirements for all four test cars (10 observations per car). These data indicate that the use of phosphorus in the tank gasoline provides reductions in rumble requirement which are essentially proportional to the concentration of phosphorus used. *It appears that phosphorus additives in concentrations up to at least 1 theory can be used effectively to combat deposit ignition and rumble in high compression engines.*

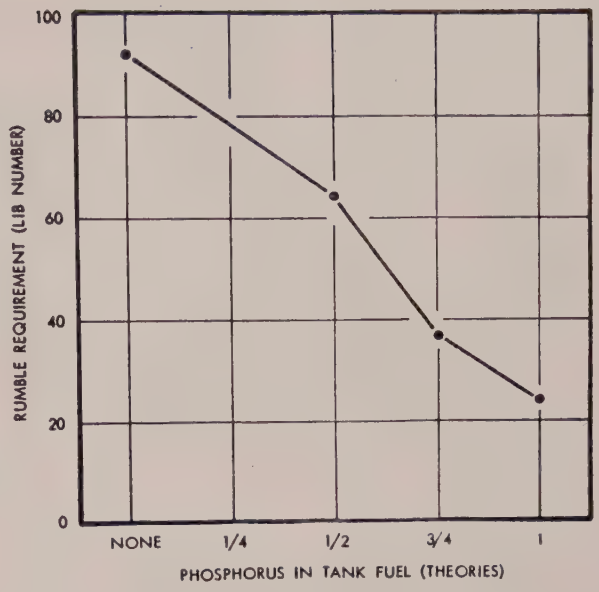


Fig. 9 — Reduction of rumble requirement when using various concentrations of phosphorus in tank fuel (average data from four cars, 12/1 compression ratio, full-throttle accelerations)

Table 3 — Residual Benefit from Phosphorus

History of Phosphorus Usage		Stabilized Rumble Requirement with No Phosphorus (LIB No.)	Requirement Test Interval (deposit mileage)
Deposit Mileage	Phosphorus		
0–3700	None	100	1500–3700
(Combustion-chamber deposits removed)			
0–6600	1 theory		
6600–9600	None	77	6900–9600
9600–17,800	½ theory		
17,800–22,500	¾ theory		
22,500–27,100	None	51	24,900–27,100

Further analysis of the data from these phosphorus tests indicates that the beneficial effects of phosphorous gasoline additives may be partially retained by the engine even though the use of phosphorus is discontinued. This is illustrated in Table 3 which includes three different values of stabilized rumble requirement for one of the test cars when using no phosphorus in the tank fuel. The three requirements are different because the history of phosphorus usage in the car was different prior to the period during which each of the requirements was determined.

The history of phosphorus usage in the tank fuel is shown in the first two columns. The stabilized rumble requirement and the mileage interval during which the requirements were determined are shown in the third and fourth columns, respectively, for periods when the car was operated with no phosphorus in the fuel.

From these results it can be observed that the use

[13] "Study of Fuel Additives to Reduce Preignition," by F. C. Burk, L. J. Test, and H. R. Jackson. Proceedings of API, Vol. 34, Section III, 1954.
[14] "Ignition Temperatures of Lead Compound-Carbon Mixtures," by G. J. Nebel and P. L. Cramer. *Industrial and Engineering Chemistry*, Vol. 47, November, 1955.

Table 4 — Advantages of Continuous Phosphorus Use

History of Phosphorus Usage		Stabilized Rumble Requirement 1 Theory of Phosphorus (LIB No.)	Requirement Test Interval (deposit mileage)
Deposit Mileage	Phosphorus		
0–6600	1 theory	4	4600–6600
0–6900	None		
6900–9600	½ theory		
9600–18,100	1 theory	48	15,900–18,100

Table 5 — Fuel Analysis Data for Deposit-Ignition-Resistance Tests

	Toluene	2,3Di-methyl-butane	Gasoline D	Gasoline E
Specific Gravity at 60 F	0.872	0.666	0.725	0.763
Tetraethyllead, ml/gal	1.0	1.0	2.90	2.91
Reid Vapor Pressure, lb	—	—	5.3	7.7
Phosphorus, theories	—	—	nil	nil
Motor Octane No.	109	104	100	93
Research Octane No.	> 120	113	111	103
Distillation, F				
initial	Boiling	Boiling	120	96
10%	point	point	178	138
50%			208	224
90%	231	136	250	327
end			314	424
Hydrocarbon Content				
Saturates, %	—	100	75	53
Olefins, %	—	—	9	9
Aromatics, %	100	—	16	38

Table 6 — Deposit-Ignition-Resistance Ratings

Fuel	Average Deposit Ignition Resistance Rating (LIB No.)	
	Car	Single Cylinder
Toluene	46	42
2,3-Dimethylbutane	73	86
Gasoline D	56	—
Gasoline E	55	—

Table 7 — Fuel Analysis Data for Deposit-Ignition-Resistance Tests

	Commercial Gasolines					
	Brand E[a]	Brand F	Brand G	Brand H	Brand I	Brand J
Specific Gravity at 60 F	0.763	0.728	0.731	0.734	0.730	0.708
Tetraethyllead, ml/gal	2.91	2.40	2.56	2.84	2.44	2.25
Reid Vapor Pressure, lb	7.7	11.2	12.0	11.9	11.3	12.7
Phosphorus, theories	nil	0.29	0.26	0.25	0.43	0.20
Motor Octane No.	93	92	92	90	92	96
Research Octane No.	103	102	101	101	101	102
Distillation, F						
initial	96	82	84	84	88	87
10%	138	121	115	117	126	108
50%	224	210	221	203	224	203
90%	327	316	350	288	322	311
end	424	398	424	378	394	384
Hydrocarbon Content						
Saturates, %	53	67	68	54	63	80
Olefins, %	9	9	11	17	15	7
Aromatics, %	38	24	21	29	22	13

[a] Same as Gasoline E, Table 5.

of phosphorus during portions of the mileage accumulation caused a residual benefit or reduction in rumble requirement during subsequent periods when no phosphorus was used. This residual benefit was more pronounced at high mileages.

An even more important point concerning the usage of phosphorus gasoline additives is that the effectiveness of a given concentration of phosphorus is reduced if phosphorus is not used continuously in the engine. The data shown in Table 4 illustrate this loss in effectiveness.

Stabilized rumble requirements are shown for two "identical"[15] cars when using 1 theory of phosphorus in the gasoline. One car had previously used fuels containing none and ½ theory of phosphorus as shown in Table 4. From these data it can be observed that the stabilized rumble requirement was considerably lower for the car which had used 1 theory of phosphorus continuously.

In summary, the foregoing data indicate that *maximum rumble reduction with phosphorus gasoline additives can be attained only through continuous use.* Intermittent use causes the phosphorus to be only partially effective. Maximum effectiveness can, of course, be restored at any time by deposit removal and subsequent continuous use.

Deposit Ignition Resistance of Fuels

The inherent deposit ignition resistance of fuels may play an important part in the control of rumble in high compression engines. Fuels having high resistance to deposit ignition will minimize the occurrence of deposit ignition and rumble even in the presence of glowing deposits.

Measurements of the relative deposit ignition resistance of several fuels were made as long ago as 1954[16, 17] in a single-cylinder laboratory engine. These studies indicate that considerable difference exists between fuels. The work described here involves rating the deposit ignition resistance of several fuels in a car on the road.

Deposit ignition resistance ratings of each test fuel were made by comparing its rumble resistance to that of various LIB reference fuel blends. This technique uses audible rumble as an indication of the presence of deposit ignition. The LIB rating of a given fuel is determined by: (1) observing the intake manifold vacuum that produces trace rumble during an acceleration when using the test fuel; and (2) determining the LIB blend that also produces trace rumble at the same manifold vacuum. The LIB number of the matching blend represents the deposit ignition rating of the test fuel — the higher the rating, the greater the deposit ignition resistance of the fuel.

Four fuels were rated in a 1958 test car equipped with 12/1 compression ratio cylinder heads. Analytical data for the fuels are shown in Table 5. Deposit ignition ratings are shown in Table 6. Each rating is the average of 10 daily ratings. Only one fuel was rated each day with approximately 200 miles of deposit accumulation between tests. Laboratory single-cylinder engine ratings[17] are also shown for toluene and 2, 3-dimethylbutane. Single-cylinder ratings were made at full throttle using a deposit injection technique.

The important conclusion from these data is that this technique can be used to separate fuels with

[15] These two cars were of the same make and model and had essentially equal rumble requirements when no phosphorus was used.
[16] "Preignition Resistance of Fuels," by B. M. Sturgis, E. N. Cantwell, W. E. Morris, and D. L. Schultz. Proceedings of API, Vol. 34, Section III, 1954.
[17] Private report, F. W. Bowditch.

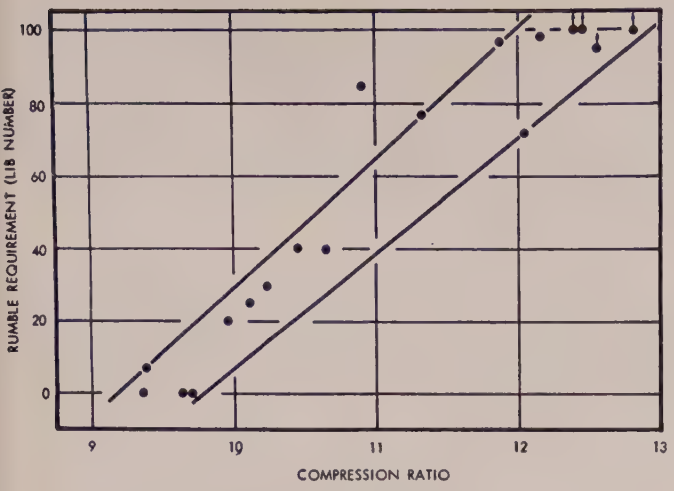


Fig. 10 — Rumble requirements of cars at various compression ratios (full-throttle accelerations)

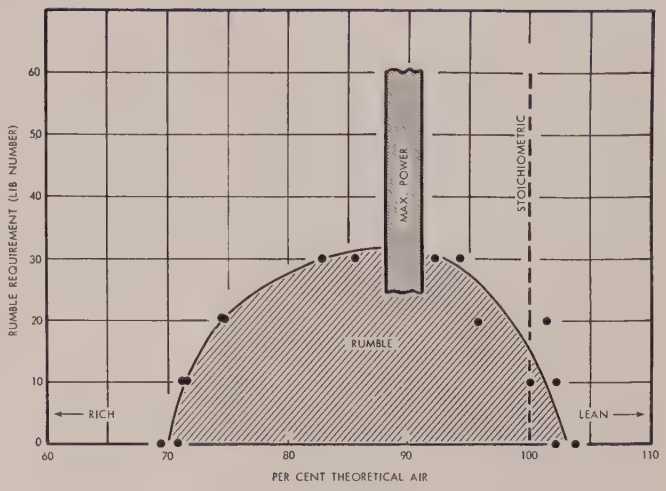


Fig. 11 — Effect of air-fuel mixtures on engine rumble requirement (12/1 compression ratio, full-throttle accelerations)

Table 8 — Deposit-Ignition-Resistance Ratings

Super-Premium Gasoline	Average Deposit Ignition Resistance Ratings (LIB No.)
Brand E	57
Brand F	57
Brand G	62
Brand H	57
Brand I	56
Brand J	61

respect to deposit ignition or rumble resistance in a car on the road. Agreement between road and laboratory ratings is fairly good.

Having established a workable technique, six commercial super-premium grade gasolines were obtained and rated for deposit ignition resistance. Analytical data for these gasolines are included in Table 7. The test car was a 1959 model equipped with a V-8 engine modified to 12.5/1 compression ratio. Fuels were rated in the same manner as described previously. Fuel ratings are presented in Table 8. Each value shown is the average of six daily ratings. The average ratings of the six gasolines are not significantly different in spite of fairly wide differences in hydrocarbon composition. It is interesting to note that Gasoline E was rated at 57 LIB numbers in these tests compared to 55 LIB numbers in the previous tests.

The ratings of these six fuels indicate that there may be very little difference among the deposit ignition resistances of present-day commercial gasolines.

Since none of the six gasolines tested have high resistance to deposit ignition, it appears that a *fertile area for further research is the development of gasolines having high deposit ignition resistance.* Single-cylinder engine studies in this laboratory indicate that fuels can be formulated which will have this quality.

Engine Operating Variables

Engine rumble, as the name implies, is not a phenomenon that can be entirely divorced from engine design. It has been observed, however, that changes in the structural design of the engine within practical limits has little effect on the occurrence of rumble. Similar observations have been made by other investigators.[7] Even though the resonant frequency of the engine parts may be altered, rumble will still occur whenever the rate of cylinder pressure rise or peak cylinder pressures are excessive. In other words, structural design changes may alter the characteristics of the rumble noise, but cannot cure the basic problem — abnormal combustion.

Certain other engine design and operating variables do have an effect on rumble through their influence on the combustion process. These include compression ratio, air/fuel mixture ratio, inlet air humidity and temperature, engine load, and engine speed.

Compression Ratio

The influence of compression ratio on the occurrence of rumble is shown in Fig. 10. Rumble requirements of several 1958, 1959, and modified high compression cars are shown as a function of measured compression ratio. The cars had accumulated 1741–30,786 miles in company transportation service. All cars were using commercial 20W motor oils but were operated with a variety of fuels. Production cars were operated on commercial premium grade gasolines and the high compression cars were operated on special high-octane, nonphosphorus fuels. Rumble requirements were determined during full-throttle accelerations from approximately 1000 to 4000 engine rpm.

With one exception the requirements fell within a fairly narrow band. The cause for this unusually high requirement was not apparent. In general, the data indicate that rumble requirement increases at the rate of approximately 35 LIB numbers per compression ratio.

Air/Fuel Ratio

Rumble is most likely to occur at or near maximum power air/fuel ratio. This conclusion was drawn from test results obtained using a 12/1 compression ratio car equipped with a fuel injection unit. Fuel injection was used for two reasons: (1) it provided uniform mixture distribution, and (2) it permitted changes in mixture ratio to be made conveniently.

Rumble requirements were determined at various mixture ratios during accelerations at wide-open

throttle in third gear at 1500–4200 rpm. The results are shown in Fig. 11. LIB number requirements are shown as a function of percent theoretical air in the mixture. Per cent theoretical air was used to express air-fuel mixtures because the individual LIB blends have different hydrogen-carbon ratios. Rumble requirement is maximum at about 88–90% theoretical air. This corresponds to maximum power for this engine as shown in the figure. A mixture of 88–90% theoretical air is equivalent to an air/fuel ratio of about 13/1 when using commercial gasoline.

Inlet Air Humidity

Humidity has a pronounced effect on the occurrence of rumble. The relationship between humidity and engine rumble requirement was studied using the dynamometer equipment and test technique described previously in Appendix I. As shown in Fig. 12, rumble requirement decreased from 60 to 5 LIB numbers as absolute humidity was increased from 21 to 134 gains of moisture per pound of dry air. This variation in humidity represents the range normally encountered during a year in the Midwest. During the tests, carburetor inlet air temperature and coolant temperature were held constant. Humidity was controlled during test observations only. These data demonstrate that humidity has a large influence on engine rumble requirement.

Inlet Air Temperature

Inlet air temperature was also found to have an effect on the occurrence of rumble. Tests were run on the engine dynamometer under constant absolute humidity and coolant temperature conditions. As shown in Fig. 13, a 60 F increase in carburetor inlet air temperature increased the engine rumble requirement by 35 LIB numbers. It should be pointed out that the effect of inlet air temperature and humidity tend to cancel each other during seasonal weather variations since humidity generally increases with ambient temperature.

Coolant Temperature

Engine coolant temperature was found to have very little effect on rumble. A range of coolant temperatures from 125 to 255 F was investigated using an engine dynamometer. Constant values of inlet air temperature and humidity were maintained. The results are shown in Fig. 14. It can be seen from the data that changes in coolant temperature produced no significant change in engine rumble requirement.

This result was not anticipated since previous road test experience had indicated an increase in rumble requirement when coolant temperature was increased. However, it must be remembered that the inlet air temperature in a car is partially dependent on the coolant temperature; therefore, when coolant temperature is increased, any corresponding increase in inlet air temperature many result in a higher rumble requirement.

Engine Load

The occurrence of rumble is more pronounced at high engine loads. The LIB rumble requirements of two 12.5/1 compression ratio cars of different make were obtained at various engine loads. These data are shown in Fig. 15. Engine load is expressed in terms of intake manifold vacuum which is an inverse function of engine load. Rumble requirements were determined during constant manifold vacuum accelerations. The requirements for both cars were progressively reduced as intake manifold vacuum was increased. This characteristic of rumble occurrence was utilized to vary engine rumble severity during the deposit ignition resistance ratings of fuels described earlier.

Engine Speed

Rumble occurs more frequently at high engine speeds. The effect of speed on rumble requirement was determined during full-throttle accelerations with a 12/1 compression ratio car. The results are shown in Fig. 16. It is evident that engine rumble requirement increases rapidly as engine speed is increased.

12/1 Compression Ratio with Proper Fuels and Lubricants

Many of the factors which affect the occurrence of engine rumble have been discussed in the preceding sections. One question remains, however, and that is: Can a high compression engine be operated without rumble when deposits are accumulated under light-duty city traffic conditions?

To answer this question, two identical cars equipped with 12/1 compression ratio V-8 engines were operated under simulated city traffic conditions. Combustion-chamber deposits were removed before the tests. The cars were driven very moderately with top speed limited to 35 mph. Average speed was 15 mph with an average of 3.6 stops per mile. Rumble requirements were determined at 200-mile intervals over a period of 4000 miles.

One test car was operated with a "good" oil (Brand F, 10W-30) and the second car was operated with a "mediocre" oil (Brand C, 20W). Both cars used a "good" fuel (Fuel A containing 1 theory of dimethylxylyl phosphate).

Rumble requirements for both test cars are shown in Fig. 17. The requirement stabilized at essentially zero LIB for the car using the 10W-30 oil and at 40–50 LIB for the car using the 20W oil. Superimposed on the figure is a band including the average LIB rumble ratings of the six commercial superpremium gasolines shown previously in Table 8.

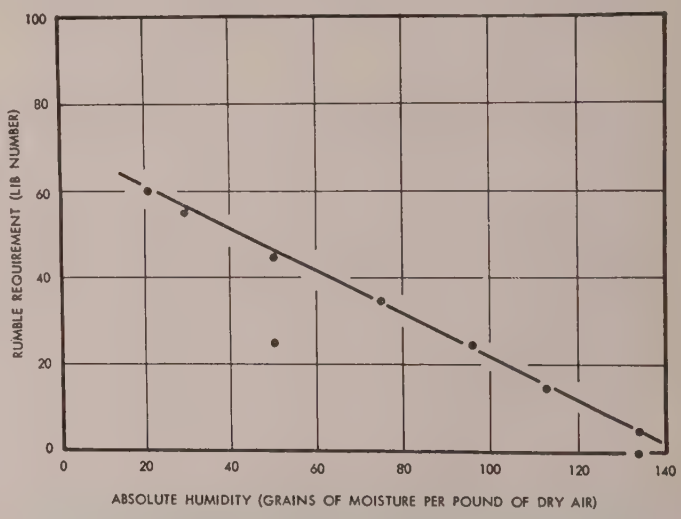


Fig. 12 — Engine rumble requirement as related to inlet air humidity (10/1 compression ratio, 3000 rpm, full throttle, engine dynamometer)

This band provides an indication of the rumble resistance of commercial-type gasolines for comparison with the car requirements. The rumble resistance of the commercial gasolines is higher than the requirements of both test cars. This demonstrates that these *12/1 compression ratio cars can be operated satisfactorily with respect to rumble under city traffic conditions if fuels and oils are properly selected.*

Summary

Engine rumble is an objectionable noise that is caused by abnormal combustion. Multiple deposit ignition causes abnormally high rates of pressure development in the cylinder which produce shock-excited resonant vibrations in the engine structure. The resulting noise is engine rumble.

Since the occurrence of rumble is dependent upon the nature of the combustion-chamber deposits, any factor which affects deposit composition will in turn affect rumble. Wide differences were found among engine rumble requirements when using 10 different commercial MS crankcase oils. In every case rumble requirements were lower when 10W-30 oils were used than when 20W oils were used.

Rumble requirements of high compression cars

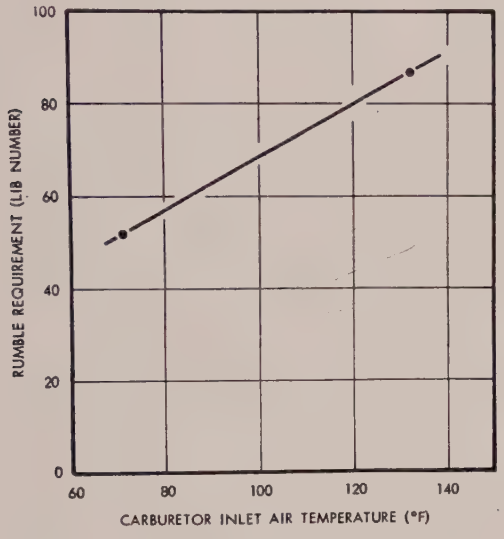


Fig. 13 — Influence of carburetor inlet air temperature on engine rumble requirement (10/1 compression ratio, 3000 rpm, full throttle, engine dynamometer)

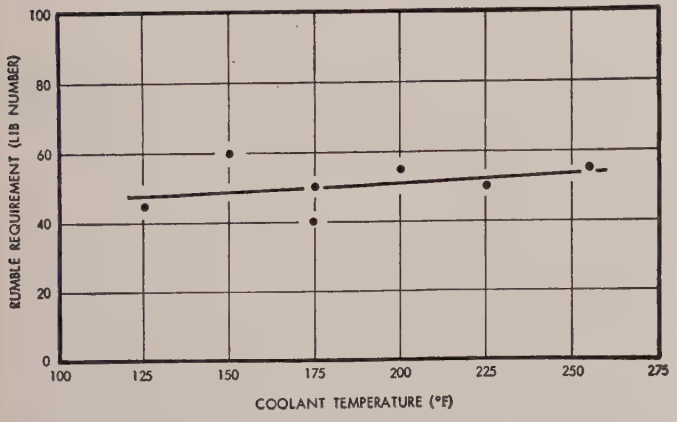


Fig. 14 — Relationship between coolant temperature and rumble requirement (10/1 compression ratio, 3000 rpm, full throttle, engine dynamometer)

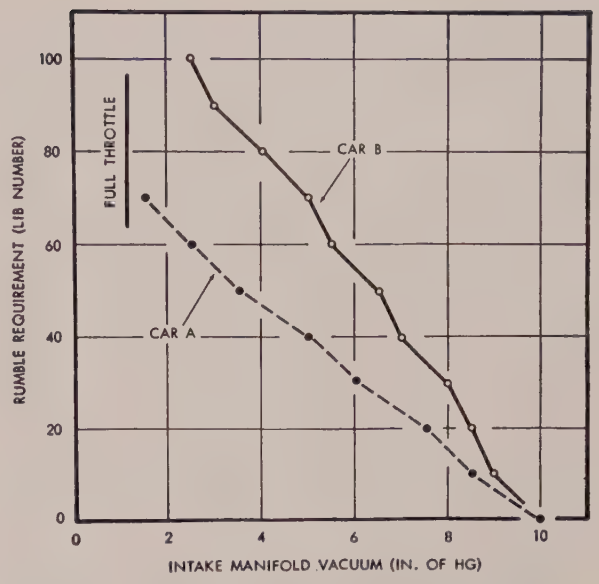


Fig. 15 — Reduction in rumble requirement as intake manifold vacuum is increased (12.5/1 compression ratio, accelerating conditions)

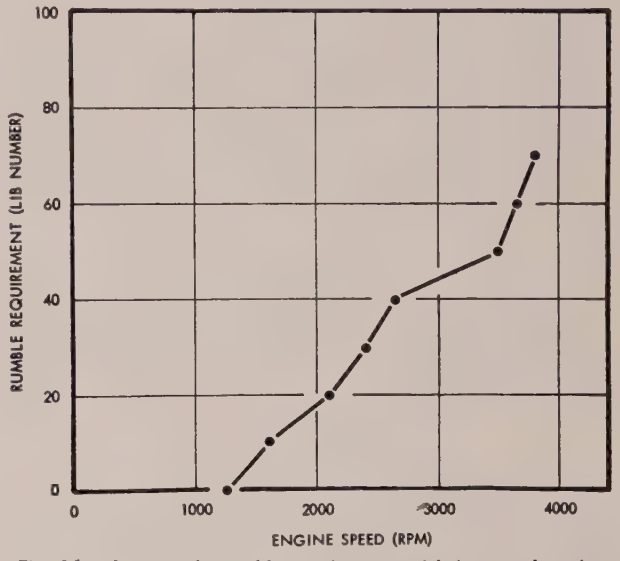


Fig. 16 — Increase in rumble requirement with increased engine speed (12/1 compression ratio, full-throttle acceleration)

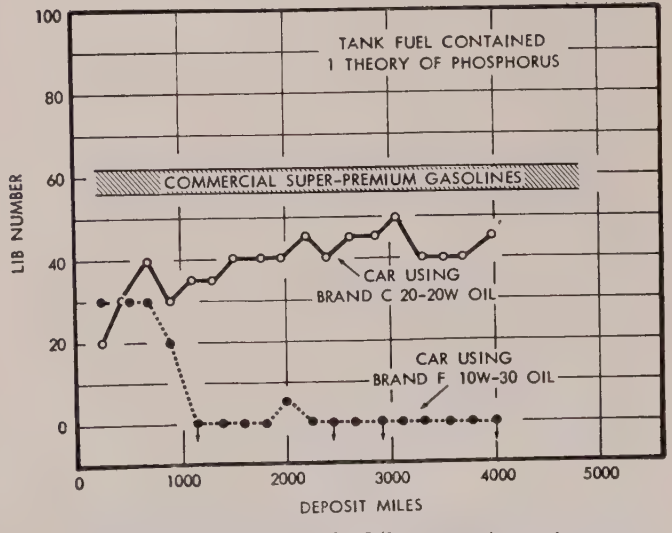


Fig. 17 — Rumble requirements of 12/1 compression ratio cars operated under city traffic conditions (full-throttle accelerations)

can also be reduced effectively through the use of high concentrations of phosphorus in the tank gasoline to modify deposit characteristics. Use of phosphorus in concentration from ½ to 1 theory provides reductions in rumble requirement which are essentially proportional to the concentration of phosphorus used. Phosphorus must be used continuously for maximum benefit.

A workable technique has been developed to rate the inherent deposit ignition resistance of fuels in a car on the road. Ratings of several fuels indicate there may be little difference among the deposit ignition resistances of present-day commercial gasolines. The development of gasolines having high deposit ignition resistance appears to be a desirable area for further research.

Certain engine design and operating variables affect the occurrence of rumble through their influence on the combustion process. The rumble requirement of an engine is increased if:

1. Compression ratio is increased.
2. Air/fuel ratio approaches best power mixtures.
3. Inlet air humidity is decreased.
4. Inlet air temperature is increased.
5. Engine load is increased.
6. Engine speed is increased.

The most important conclusion from these studies is that higher compression ratios and correspondingly higher engine efficiencies can be reached without objectionable rumble if oils and fuels are carefully selected. It has been demonstrated that cars can be operated satisfactorily at compression ratios as high as 12/1 even under light-duty driving conditions.

Acknowledgment

The authors wish to express appreciation to members of the Fuels and Lubricants Department for their assistance in collecting data and in preparing the manuscript.

APPENDIX I

Dynamometer Test Procedure

Equilibrium combustion-chamber deposits were accumulated using two 1959 production V-8 engines. The engines were operated on a light-duty deposit buildup schedule in which engine speed was held constant at 2000 rpm. The load was varied intermittently as indicated below.

Load	Manifold Vacuum	Duration
12 to 15 hp	20 to 21 in. Hg	30 min.
No load	26 in. Hg (approx.)	20 sec
120 to 130 hp	1 in. Hg (approx.)	20 sec

This load sequence was repeated continuously for a total of 200 hr. The coolant temperature during deposit accumulation was 175 ± 5 F, and the carburetor inlet air temperature was 75 ± 5 F. No provisions were made for oil cooling, so oil sump temperatures ranged from 195 to 225 F depending on the oil used. Manufacturers' recommended spark advance was used throughout the test. Except for special tests, the carburetor inlet air humidity was determined by ambient conditions.

Engine rumble requirements were determined periodically at 3000 rpm — full throttle with the coolant temperature at 200 F. A minimum of 7 hr of deposit accumulation preceded each rumble requirement observation. The individual requirements obtained between 100 and 200 test hr were averaged to provide a stabilized rumble requirement for each test.

APPENDIX II

Pan Evaporation Test

The pan evaporation test consists of the evaporation of samples of test oils at 550 F in a circulating-air oven. Each oil sample, weighing approximately 2 grams, is placed in a tin-plated dish, 1¾ in. in diameter. The sample is then placed in an electric oven which has been preheated at 450 F. Air is circulated through the oven at the rate of approximately 60 cfm. After the oil samples have been subjected to this temperature for one hour, the oven temperature is increased to 550 F and maintained at that value for 1 hr. The samples are then removed from the oven, allowed to cool, and weighed. The per cent residue for each sample is calculated from the weight data. Duplicate tests are conducted on each test oil to obtain an average residue value.

ORAL DISCUSSION

Reported by R. D. Jolly
Standard Oil Co. of British Columbia Ltd.

R. E. Steinke, California Research Corp.: The per cent of residue in the pan evaporator test showed a relationship to engine rumble requirements. Has General Motors tried ashless detergent oils in their engine test for rumble determination, and if so, what were the results?

A: Yes, they have been tried and the results were good, although we doubt that the detergent had any large effect.

Dr. E. S. Starkman, University of California: Has actual engine damage been found in the field which could be related to rumble? Can clean engines be made to rumble?

A: No incidence of engine damage solely related to engine rumble has yet been found; however, damage has been observed when nonknocking surface ignition occurred. Knock and rumble together can fail engines.

Clean engines can be forced to produce a noise similar to rumble, particularly in high compression ratio engines, (for example, 12/1), or with highly advanced spark timing. This noise is sometimes referred to as "thud," although General Motors prefer to refer to this as engine roughness.

Gilbert Way, Ethyl Corp.: How significant is rumble and how prevalent? Ethyl's experience has been that even with compression ratios of 9.5/1, rumble can be encountered. It is also apparent that fuel composition can affect tendency to rumble. In particular, heavier aromatic fractions induce rumble, causing deposits. The resistance to deposit-induced ignition of light olefins seems good. There was some evidence that the use of phosphorus improves the resistance of an engine to rumble.

A: General Motors' Car Division have had some complaints of rumbling, but have not been able to determine that the complaint originated from rumble without knock also being present. Knock disturbs chamber deposits, and under some circumstances, this can increase rumble.

Harold F. Kley, Shell Oil Co.: Did the superpremium gasolines in the field studies contain phosphorus?

A: Some did, some didn't.

Mr. Kley: If some fuels contained phosphorus and some didn't, why wasn't the band of commercial superpremium gasolines shown on Fig. 17 wider?

A: The band shown on Fig. 17 represents the deposit ignition resistance or rumble rating of several commercial super-premium gasolines. We do not think that phosphorus affects the deposit ignition resistance of a fuel. However, phosphorus does affect the combustion-chamber deposits formed when it is used and, therefore, reduces the rumble requirement of the engine. The distinction is between *fuel ratings* and *engine requirements*.

ways of improving

TAKE-OFF AND LANDING

F. W. Kolk
American Airlines, Inc.

SOME POSSIBILITIES for shortening the field length requirements of present-day jet aircraft are:

- Install leading-edge, high-lift devices which are retrofitable to present-day aircraft.
- Retrofit — or purchase new — aircraft powered by turbofan engines. These have an inherently higher take-off thrust to cruise thrust ratio than the jets, which vastly improves the take-off acceleration.
- Use boundary-layer control actuated by turbine discharge gas for immediate consideration in new aircraft engines.
- Use direct-lift jet engines. These will improve the block speed characteristics of the aircraft and also give vertical take-off and landing capabilities.

This paper discusses the advantages of each of these possibilities. The author also describes the problem of airport location within a city, and its effect of total travel time.*

FIELD LENGTH for operation of turbine-powered aircraft has been a major problem in airline operation. The practical result of this problem has been the construction, on an emergency basis, of a number of runway extensions at various airports around the country for the operation of the big jets. I think that everyone is concerned with whether this round of extensions will be the last one — or whether there will be a requirement, several years hence, for an even greater round of extensions. What is the practical limit on runways and have we passed the point of diminishing returns?

There's an old saying about the road to hell being paved with the best of intentions. That is very apt at this juncture. The present difficulties we find ourselve in, as an industry, are not all our own fault; we have had a lot of help in muddying up the waters. I'm now going to throw another glob of mud at some of my old friends in the Civil Service, knowing full well that I'd better duck right fast for the return volley.

I think that one point must be made very clear — Airlines and aircraft manufacturers did not trip off the present situation with malice aforethought. Estimates were made — carefully made — of the available runway lengths and these available lengths were matched to the estimated aircraft requirements by the airlines before these aircraft were ordered. Nor, in all due fairness, did the manufacturers of the aircraft deliberately underestimate the requirements for runway length. What did happen was that the industry was faced with a set of retroactive rule changes (made after the purchase commitments for the aircraft) which progressively raised the standards of required performance. Without any qualification as to justification, let me

* Paper presented at SAE National Aeronautic Meeting, New York, April 3, 1959.

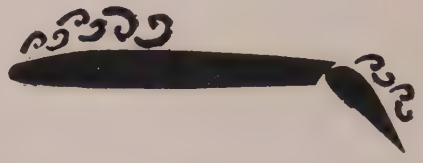


Fig. 1 — Thin airfoil with flap

Fig. 2 — Leading edge slats

repeat, the industry has been faced with a series of retroactive and progressively more conservative changes in rules and interpretations of rules which wiped out all of the planning relative to the application of aircraft to specific runways. This condition was indirect violation of the established order of business in certification of aircraft, which provided that a certification must be issued under the rules existing at the date of the filing of the application for a type certificate.

This is a rather serious charge to level against a responsible agency of the Federal Government. For those who would question my presumption, I would refer to the Preamble and the explanatory material issued in the Federal Register in the notice of rule-making for SR-422, SR-422A, and lately SR-422-B.

Just so that I do not give the impression that the Bureau of Safety Regulation is the only one guilty of changing horses in midstream, I would like to point out that there is a remarkable degree of inconsistency between the Administrator's interpretation of essentially the same language in CAR 04B and SR-422.

Without going into the matter of whether these changes were justified, let me point out that they occurred; that the power used to accomplish them was reserved for use only under extreme conditions of safety emergency and that, on such a basis, no adequate justification has ever been offered the American people.

My reasons for so thoroughly ventilating these ideas is to point out to the public that the airlines are not irresponsible people — that they did not purchase aircraft which would not fit into the existing airports. The situation that now faces us is that first, the operating rules are somewhat more severe than were anticipated, that the performance of the aircraft under these rules is a little disappointing, and that the problem of air traffic control has turned into a vastly more difficult problem than anyone thought it would be (here is a subject for an entirely new sermon). We are faced with the fact that we don't have enough runway length around the country to operate the jets properly. What do we do about it? First, we either restrict loads, or we do not offer services which the aircraft are potentially capable of performing because of inability to operate safely at the limited weights allowed. Second, we ask for the elongation of many important runways. This phase takes a little longer and in some cases cannot be accomplished at all without abandoning some existing airports. Third — and somewhat down the line — we attempt to modify the aircraft to improve its take-off performance. Fourth, as a long range project, we look at the next round of aircraft and tell ourselves not to get into this trap again — let's do something different!

I am going to deal with the latter several phases of the problem; that is, modification to the aircraft, either on a new or retrofitted basis, in order to shorten the field length requirements.

The problem of airport performance is not a new one. It is merely the problem of the distance required to accelerate to or decelerate from the minimum safe flight speed. The mathematics are classical in nature and, in fact, are the very ones Newton was concerned with when he invented the calculus. For our purposes we can write, in a very much simplified form:

$$S = \int_o^s ds = \int_o^v \frac{V\,dV}{a} \qquad (1)$$

or, if we pretend that the acceleration is constant:

$$s = \frac{V^2}{2a} \qquad (2)$$

We see that, generally speaking, the distance will vary as the square of the speed or inversely with the acceleration. Now, the safe speed is some percentage above the lowest speed at which the aircraft will remain airborne under control. In short, the stall speed, neglecting atmospheric variables, is a function of the wing loading divided by the lift coefficient.

$$v_s = \sqrt{\frac{2W}{\rho S C_L}} \qquad (3)$$

For take-off, the acceleration is a function of the thrust-to-weight ratio. The excursion into freshman physics is made not to develop anything profound but to remind you that you can't get something for nothing in field length.

Leading-Edge, High-Lift Devices

Over the years we have seen a gradual increase in the usable Cl_{max} through the use of better flaps, until such time as flight speeds entered the compressibility region. Since that time, it is sad, but all too true, that anything done to improve the high-speed characteristics of a wing reduces its maximum lift. From a take-off and landing viewpoint, our wings are decidedly less effective in the newer, faster aircraft. It has been noted, moreover, that with the thin wings required for good, high subsonic operation, the resulting sharp leading edges cause premature stalling of the forward upper surface. (See Fig. 1.) This can be cured by camber, but such camber tends to ruin the high-speed ability of the air-

foil. The answer is the use of leading edge devices, such as dropped lead edge, Kruger flaps, slats, or slots. (See Figs. 2 and 3) Curiously enough, such devices are the rule rather than the exception on military aircraft, but until recently their use on commercial vehicles was not even considered. There is now some indication that the well-dressed transport will sport her slats, and will thereby benefit by increased lift, decreased drag, and ability to fly safely very much closer to the stall than heretofore has been possible. Already the Boeing 707-120 series is equipped with a short section of Kruger flap just inboard of the outboard strut, and her younger sisters of the 707-020 series will be far more generously endowed. The Douglas DC-8 is flying with a slot arrangement and the Convair 600, the big sister of the 880, will sport slats on most of her leading edge. These things shorten the take-off and landing distance.

Turbofan Engines

In dealing with this problem, it is wise to note that either the take-off or the landing distance can be critical, depending on the distance to be flown and the characteristics of the individual aircraft model dealt with. Generally, it is the take-off distance that limits the operation aircraft. Therefore, we must look particularly for ways and means of attacking this phase of the problem — keeping in mind that landing might become critical if take-off is improved very much. There is one possible retrofit that reduces the take-off distance without affecting the landing distance which will be becoming much more popular in the near future. That is the fitting of the turbofan-type engine in place of the straight turbojet.

Why will the turbofan help the take-off distance? By two very simple and important ways. If we measure the size of the engine in terms of the cruise thrust at altitude, then a turbofan of equal size to a turbojet will have more take-off thrust. It will, therefore, have a greater take-off acceleration. The ratio of this increase in thrust for take-off is the order of 20% in the engines currently being offered. We can expect that the distance required to take-off at a given gross weight will be reduced by a factor of the order of 20%. At the same time, the turbofan is a considerably more efficient powerplant in cruise than the straight turbojet, and the specific fuel consumptions will, therefore, be somewhat lower. Because of the lowered fuel consumption, the take-off weight will be reduced to fly the same distance; and in aircraft of this type, for every per cent that the weight is reduced, the take-off distance will be reduced by about 2%. Thus, for fairly long ranges, the turbofan aircraft will require from 65–75% as much runway as the turbojet aircraft to do the same transportation job.

This is not the only reason for installing the turbofan engine and it is profoundly hoped that the majority of our aircraft will be so equipped in the near future. If the builders of engines had listened to the pleas of the operators sooner and not waited until competition forced them into the turbofans on a "crash" basis, many of the runway extensions now planned would not have had to be undertaken. So much for spilled milk.

In our opinion, this gets to the end of the road for the present aircraft on a retrofit basis. Anything beyond the installation of turbofan engines and leading-edge high-lift devices is difficult to conceive as a retrofit, without prohibitive costs. Anything else will require at least new production machines.

Suppose now, that we have gone as far as we can down the road of retrofit, and we have substantially reduced — that is from 10–30% — the take-off distances, and have made some considerable progress with the landing distance. And we have gone along with our program of fixing up the worst of the dogs of runways. Everything fits. The situation is then like the state of affairs in piston-engine operation just prior to the dawn of the Jet Age.

The equilibrium could be disturbed by several developments. First, in a desire to create a new market for aircraft, some builders may be led to invade the very short field market now dominated by the propeller-driven aircraft with small jets. This means the design of jet aircraft with an honest field length capability of 4500 ft. Such performance is clearly beyond the capabilities of jet transports as presently conceived. As it stands at present, the wing loadings are far too high for the low flying speeds for landing and take-off required for such short operations. And, if sufficient wing area were to be used, the resulting riding comfort in mild, low level turbulence would pretty well debar the use of this class of vehicle. What we need is to be able to raise the wing loading and at the same time lower the stall speed.

The answer to this type of vehicle seems to be boundary layer control for the development of lift.

Boundary Layer Control

BLC, as it is referred to, can be used either to increase the lift of a surface or to decrease the drag of a surface. In either case, the desired result is achieved by reducing the tendency of the flow to separate as it encounters adverse pressure gradients. This is done either by sucking away the spent boundary layer or by revitalizing it by the introduction of high velocity air at or near the point of separation. If the system is to be successful, it is necessary that the power required to eliminate the separation or to

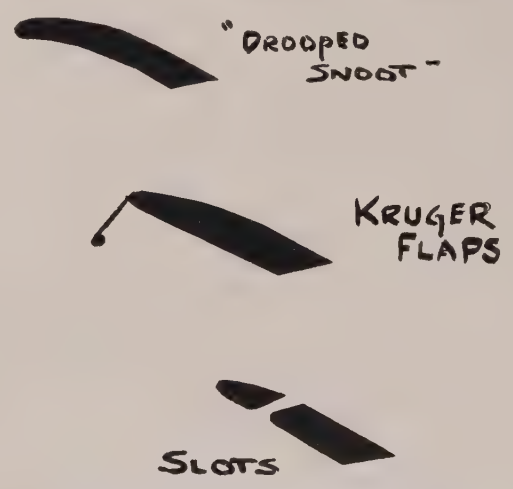


Fig. 3 — Leading edge high-lift devices

Fig. 4 — BLC system being used to control separation occurring on rear upper surface of highly cambered airfoil

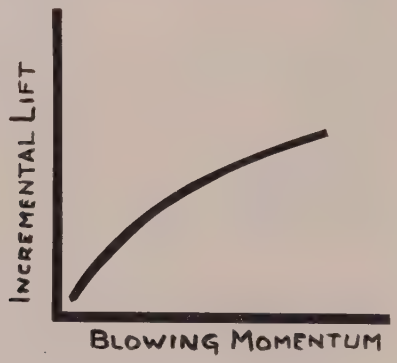


Fig. 5 — Curve showing effectiveness of BLC system

drive the system, be less than the power saved in operating the equivalent vehicle without the BLC. As this is sometimes very difficult to prove, perhaps we should say that the BLC system is successful if the vehicle built and operated with it is a better vehicle than one not so equipped.

Taking first things first, as it were, we can say that any system to be used on a commercial aircraft must be reduced to reasonably sound practice before it is considered for such use. This fairly well limits us to blowing systems designed to increase lift. All such systems blow air out of slots to increase the lift of the wing.

The mechanism is shown schematically in Fig. 4. In this case, the BLC system is being used to control the separation which occurs on the rear upper surface of a highly cambered airfoil. By so doing, the lift of the airfoil is considerably increased, and the drag is reduced — at least the profile drag is reduced.

The induced drag, if we have a finite airfoil, may get to a pretty high value if the lift is sufficiently increased.

How do we measure the effectiveness of the BLC system? The general idea is shown in Fig. 5. It can be seen from the data that the first little bit of air does a lot of good but the incremental effectiveness falls rather rapidly as greater and greater force is applied to the system.

Tests show that rather large lift coefficients are available to the aircraft designer by the use of such arrangements. Why do we not dash out and install them on our aircraft forthwith?

If we were in trouble on the landing requirement and just didn't have to worry about the take-off, I think that the answer would be fairly clear — we would have BLC on almost everything that has wings. In this condition, the engine is not required for acceleration except on a pullout, and is available as a source of power to drive the system. There is no net loss involved in this drive requirement. On the take-off, however, if the engine is used as the means of acceleration of the vehicle up to flying speed, any power used by the BLC is going to come out of acceleration, unless some valving arrangements are made for cutting in BLC at breakground. Such valving may be beyond our present state of the art, inasmuch as it affects aircraft control systems, engine control systems, and pilot technique. What a time to have an engine failure!

Actually, there has been some progress made with BLC for take-off. When people first started to build systems of this sort for jet aircraft, there was no net gain in take-off distance. Now you can build a system, and some have been built, where the take-off distance is actually somewhat shorter with the BLC system on than off. The reason for this rather poor luck with the BLC installations is that they have depended on the bleeding of compressors to get the high-pressure air for the blowing system. When this is done 2–3 lb of thrust will be lost from the engine for every pound of thrust that is applied to the BLC.

Why has BLC not lived up to its exceptations in commercial vehicles? Purely because the state of the art has not permitted it. There are many available references to the physics of BLC and there have been many specific tests on complete aircraft — both in flight and in the wind tunnel — which show the underlying promise of this system. Yet we have not seen anything spectacular in the way of improvement in aircraft efficiency as a result of the application of this system. There must be reasons.

To start with, most serious applications of BLC have been on jet aircraft and in each case, the air for the blowing system has been made available through the process of compressor bleed. The application of this system to a jet engine is very injurious to the resulting thrust of the engine. If we keep in mind that the BLC blowing system uses a jet of air blown out over the flaps to achieve its results, it would seem quite obvious that this jet of air is available — at least during the ground acceleration for thrusting the aircraft — and that when it is doing its work as a lift augmentor it has to overcome the additional induced drag of the regime plus some entrainment losses. Clearly, it is not going to waste. If we take a pound of thrust away from the jet and reapply it to the trail edge of the wing, we should be paying Paul even if we are robbing Peter. Of course, because of duct losses, poor nozzle coefficients, and temperature losses, we cannot hope to get the same energy recovery out of the blowing slots that we got from the main jet — but it should let us have ample room to work out a system.

However, that isn't the way these things get built.

When the compressor is bled to get the blowing air, for every pound of momentum we can theoretically produce in the blowing jet, we have robbed ourselves of 2–3 lb of thrust from the engine. If we add system losses to that, it is quite evident that any major reduction of stall speed will entail a very major reduction in acceleration. Thus, we come to a basic conclusion, namely, that the BLC system must be so integrated with the powerplant that the most efficient compromise of arrangements is used.

But we cannot stop here, as we have not yet come to grips with the knotty part of the problem. The BLC system is an airworthiness item to the transport aircraft. It cannot be subject to loss of performance or control due to failure of a powerplant. And if we get our air from a powerplant, the system must be designed to function with a powerplant suddenly made inoperative at any point during the take-off and at any speed which will likely be encountered during the take-off. In short, it must comply with T-category. This means that if the arrangements are such as to change the BLC jet force as a result of the failure of one or more sources of blowing power, this change must be symmetrical and the operating speeds of the aircraft should be selected so that there is ample margin above stall, even at the reduced flow. It would appear, then, that a multiple source of blowing air must be used.

To me, this says that the real way to achieve BLC is to duct turbine discharge gases through the flaps blowing system, because this system produces the only minimum loss source of thrust for BLC. It further appears that the turbofan engine may be the best configuration of basic powerplant to supply this hot gas, becauses its gasses are already somewhat cooler than those of the straight turbojet by virtue of the energy removed for fan drive. It would also appear that some form of variable-area exhaust nozzle is especially applicable to such a system, to make up for the gasses being diverted to the flaps. Simple flapper-type devices can be used to account for system on or off and a second, smaller set of flaps can be used to compensate for the loss of an engine and can be controlled automatically. Thus, the engine can be made to operate on a single basic operating curve, does not have matching problems as a result of system operation, and produces BLC momentum efficiently. Its two disadvantages are the temperature of the gasses in the pipe and the fact that the flaps will be black on such aircraft, regardless of the wishes of the exterior decor people.

There is a second benefit to be gained by the use of BLC during take-off. This arises out of the fact that we are now faced with an all engine take-off requirement. Up until now, the engine out take-off condition has been critical and the designers of aircraft have done all they could to bolster up this performance. One of the things done was the reduction of the wing incidence to as low a value as possible to make the brakes more effective. This feature, in connection with anti-skid devices on the brakes, have made the braking action of the aircraft in a rejected take-off condition unbelievably good. At the same time, it has caused a real problem in rotation — especially in the all-engine take-off condition. Regardless of what you might hear about the rotation speed and the speed gained during rotation, our last generation of aircraft didn't have this problem because they didn't have to rotate to take off. The problem with the B-377, for instance, is to keep it from ballooning up off the ground too soon, a quite different condition to today's aircraft.

Actually, what is wanted is the ability, in effect, to vary the incidence of the wing, whether a take-off is being completed or rejected. And BLC does just this for us. When the engines are idled, the blowing stops over the flaps and the lift coefficient drops abruptly. Or, if you keep going, you get the effect of the higher incidence and lower rotation requirements at take-off. This could be a way of living with SR-422 and liking it!

VTOL Aircraft

Everything that was said about the conventional jet aircraft could said about the supersonic transport, only more so. The extremely thin wings required for flight faster than sound, with their extremely sharp lead edges and radical platforms make the achievable life coefficients very low and the probable approach drag coefficients very high. I feel quite sure that a very sophisticated BLC system will be at the heart of the concept of a practical aircraft of this type. There is, of course, the ultimate in short runways — the VTOL aircraft. And here we find many, rather than too few, configurations that are accepted. As I am looking at this problem from the viewpoint of one who is engaged in mass transportation, you will forgive me if I lack enthusiasm for a vehicle merely because it can hover, or land or take-off short. It must also be a useful vehicle — not only in an absolute sense, but also competitively to other alternative vehicles. The present-day crop of VTOL, including the helicopters, have one element in common — they are useful only under some limited circumstances when the VTOL ability transcends all other requirements. However, the places in the world where nothing but a VTOL will serve the purposes are in the minority.

The concept of a VTOL aircraft operating from the centers of metropolitan areas to the centers of the metropolitan areas has been taken for granted by such a large proportion of the people concerned with aviation for so many years that I feel I cannot reject this concept without offering a valid, binding explanation for this rejection. We can visualize, however, the typical journey and the effect upon this journey by several different assumptions concerning the location of airports and the ability of aircraft to operate therefrom. Let us postulate that the typical intercity journey consists of three parts.

First, a journey from the domicile of origin, either the office or the home of the traveler, to a transfer point from which high-speed, intercity transportation operates. Then the high-speed ride to the transfer point to the domicile or premises of distance, be it the home or the office to be visited. We can simulate any transportation system by varying the characteristics of each one of these links and, for our purposes, we can investigate the first link which we can assume to be a mirror image of the third link. The popular theory is that if the airport — or heliport — were located in the exact center of the downtown area, that the link will become either smaller or disappear completely.

The classical situation which eliminates the airport postulates that everyone starts from the exact

center of the urban population at one point and ends up in the exact center of the urban population in another point and that it is possible to build an airport or some other catcher's mit and sling-shot combination at this exact center where all these people are clustered. This supposes that this magic, small spot is going to be pretty heavily occupied by people, offices, dwellings, airplanes, and various and sundry other incompatible impedimenta. We arrive at this theory through analogy to the railroads who have succeeded in locating their railroad stations very close to the downtown centers of population. At this point, I would like to observe that the airport analogy is tremendously oversimplified in various people's minds. To start with, the reason the railroad stations are at these centers is because the centers grew up around them and if you trace the urban development of virtually any city you want to take at random, except Los Angeles, you will find that this concentration close to the railroad station grew up after the railroad was built. Los Angeles is the classical example of the exception to the rule in that it antidates the general trend towards decentralization.

There isn't any center of Los Angeles because of the dependency upon private automobiles for general transportation. I think it is safe to say that the usual situation in an intercity trip presupposes that this initial link from the origination premises to the transfer point — airport, heliport, or what have you — shall be made in either a private automobile or a taxi in the general case, with a secondary tendency toward a bus or limousine or some similar vehicle. This is going to be the case more and more as our cities spread out. It must be true for better than half of the conditions because most people who travel on an intercity trip, either start or finish each individual leg at home. It is not rational to postulate that a man will leave his office in New York City, go to Boston for an appointment, and then return to his office, as a habit. He will either leave his office, go to Boston, and then return home — or go to Boston directly from home or, if it is going to be an all-day junket, he will start at home and end at home.

It is also difficult to visualize any transfer point just within walking distance of the passenger. Cities are just too big for this. No matter where you put a heliport on Manhattan Island, the majority of the passengers to and from this heliport will have to use auxiliary ground transportation in order to complete their journey satisfactorily and evolution tells us that the cities will spread out.

We have now two choices to make:

1. To consider one central transfer point for all of the operation between the two large cities and the other is to visualize a multitude of transfer points in each city to minimize the ground transportation problem at the ends. Referring to Fig. 6, I have sketched an area which contains a homogeneous distribution of population from which trips can originate or to which trips can be directed and I have sketched an airport of varying locations from the center of this area to well out in the sides. If the airport is a long distance from the center of the city, in terms of the city boundaries, the average distance from any point to the airport will be approximately the average distance from the center of the city to the airport. But if we move the airport to the center of the city, the average distance will not become zero because we are generating traffic from the whole area of the city rather than the exact centroid. In fact, if the traffic generating potential were evenly distributed by area, the average distance to an airport located in exactly the center of this area would be two-thirds of the radius of the area. So, in the case of one airport or transfer point, we cannot eliminate the ground time unless all of the travel is to and from the airport and we cannot logically assume that this situation is a general one.

Suppose now we try to carry this situation by dividing this city area up into zones, and we will arbitrarily divide our squared city into four equal pieces as in Fig. 7. If we locate a helicopter or airport in the exact center of each zone, we have not eliminated the ground link at the beginning and end of the flight; we have merely made the average half as long and we have not, in any sense of the word, seriously reduced the travel time in this ground rule. You do not make a taxi ride half as long by making the distance to be traveled one half of what it was. You perhaps only make it two-thirds as long. But what have we accomplished? Fig. 8 shows two cities — each divided into four parts and we can do a little bit of reflective figuring to see what problem we are up against. Suppose now these cities generate 100 passengers per hour between the two and we serve these cities with 100 passenger airplanes operating at a 50% load factor. This means that we have to operate 200 seats per hour between the two cities and this means one flight every 30 min. Under these conditions, the average length of time it takes to make a trip originate at a random time period is the

Fig. 6 — Airport location versus traveling distance from town to airport

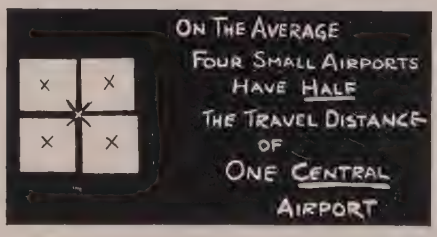


Fig. 7 — Traveling distance to decentralized airports

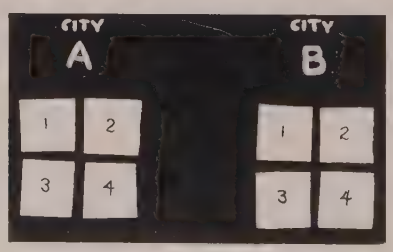


Fig. 8 — Travel between two cities with decentralized airports

trip time plus 15 min or half of the average interval between flights. And if we have a 15-min taxi ride at each end, this is a 1-hr flight. We have a total average time of 1¾ hr to make the trip. Now, let us look at our four scattered heliport per city situation. We have perhaps reduced our average taxi time by this maneuver to 10 min at each end or a total of 20 min and we have a 1 hr and 20 min total running time plus some waiting time to accommodate the schedule frequency. However, instead of serving two pairs of airports, we now have to serve 16 pairs of airports because anyone who happens to be located in city division A-1 has an equal probability of wanting to go to B-1, B-2, B-3, and B-4. There are other people in A-2, A-3, and A-4, all of whom have the same desires. Therefore, we can serve only an individual pair of airports one-sixteenth as often as we could before, or the average period between flights would be 8 hr. This means now that in order to overcome this frequency problem, a traveler either has to accept the transportation from airport pairs which are in his most convenient location or the operators must reduce the capacity of his flying machines from 100 seats to something of the order of six seats to make the same schedule frequency. It is quite obvious that neither of these solutions is to the public interest.

So, we can conclude from all of this abstract reasoning that it is best to have only one airport per city and that this airport can equally well serve the population of this city from somewhere on its outskirts as well as its exact center. The only disadvantage to the one airport per city situation is when this airport is so far from the city limits that excessive running time results for everyone concerned with making a trip.

If any of you doubt the validity of this reasoning, try this as a stunt. Think back to the last time you visited Los Angeles and then try to rationalize where, in Los Angeles, you would put the airport if you were to make it most convenient for this last trip. Then find someone else at random, who has been to New York recently and ask him how much his trip would have been improved had the airport been located where you decided it might best be for your purposes. I would venture to predict, with at least a 50-50 chance, that the total running time for both your improved trip and his resulting trip from airport to destination would not be any less than the sum of both of your running times on your last trips from the present location of the airport.

We can say, as a result of this diversion, that what we really needed out of airports is a location somewhere within a city rather than 30 miles out in the country and that to the general, broad, run-of-the mill traveler, it doesn't matter much where, in the city, this airport is because what will accommodate one man will disaccommodate the next but that everyone realizes a benefit from a single airport.

Now, quite obviously, we do not need to have VTOL aircraft in order to get airports somewhere within the city areas but in the same breath, we must realize that we can't have earthshaking monsters, requiring 3 miles runway for this sort of service either. Is there any benefit then from vertical rising capability within this framework?

It is becoming increasingly apparent that the major limits to high-speed operation between cities of moderate distances in operation is the time the aircraft spends on the ground nonproductively, taxiing, awaiting take-off, awaiting gates, and the like. These unproductive minutes are becoming an increasing economic burden to the aircraft operator and, therefore, to the general public. The time is caused by two factors; the necessity for taxiing long distances, which increases as the runway length increases and the interferences which many aircraft on the ground create for each other by clogging taxiways, blocking take-off positions, and generally interfering with each other. By and large VTOL aircraft, at least such VTOL aircraft as can take off from their landing positions, will eliminate this time involved. That this time is worth salvaging may be more apparent if you realize that in October, 1958, certain flights of my company averaged 25 min on the ground, in taxiing, awaiting clearances, running up, and awaiting gates. This degree of confusion appears to be getting worse rapidly as the number of aircraft movements on a given airport increases. It is quite obvious that even at best, traffic control on the ground can only hold this congestion to perhaps present levels. A VTOL aircraft can eliminate some 15 min average ground time from the present operation when operated from present-day airports and obviously can eliminate the need for extremely long airports so burdensome to communities. If such a machine can be built so as to decrease the net travel time between cities and at the same time be entirely competitive or economically superior to the present style of aircraft, it will stand on its own feet and be widely accepted.

One configuration which has tremendous promise in this respect is the so-called "hover" jet aircraft configuration, which derives its lift for take-off and landing from a multitude of vertically positioned jet engines but which cruises as a conventional aircraft with very much reduced wing and cruise powerplant dimensions. This kind of machine appears eventually to be competitive for even fairly respectable ranges, depending upon the achievement of extremely low weight-to-thrust ratios and extremely small volume-to-thrust ratios for the lifting powerplants. Certain of our engine companies have made a start towards producing such a powerplant and while they are not quite to the point of economical feasibility, they are progressing enough so that we can afford to entertain this possibility.

Conclusions

To summarize then — I have made the following specific recommendations to improve the take-off and landing characteristics of aircraft:

1. The installation of lead edge, high-lift devices which are retrofitable to present-day aircraft.
2. The retrofit or new purchase of aircraft powered by turbofan type of engines which have an inherently higher take-off thrust to cruise thrust ratio than the jets and, therefore, vastly improve the take-off acceleration.
3. For the future, the use of boundary-layer control actuated by turbine discharge gas for immediate consideration in new aircraft designs.
4. The use of direct-lift jet engines to improve the block speed, characteristic of the aircraft, at the same time giving vertical take-off and landing capability.

Properties of Silicone Rubber for

INVESTIGATION of silicone rubber compounds for static seal applications has demonstrated the importance of tests to evaluate these materials for specific properties and at environmental temperatures which simulate anticipated service requirements. Some relationships have been established between these properties and compound composition; they are presented as a guide for the selection of suitable materials for specific seal applications.*

SINCE RUBBER products cannot be adequately controlled for procurement purposes by composition alone, specifications are generally based on performance characteristics and define these materials in terms of physical property capabilities. Many specifications attempt to control a material for a variety of shapes, forms, and applications with a single set of conventional property requirements which may be completely unrelated to the peculiar functional requirements demanded of the material in specialized applications. The result is that these specifications fail to define accurately the specific properties necessary for the successful operation of such parts as seals. This has been the situation for existing silicone rubber specifications.

Increased operational temperature requirements which led to the adoption of silicone rubber for seal applications have widened the difference between room temperature properties normally measured in specification testing and the realistic properties which materials exhibit under service conditions. This factor is separate from heat stability of such materials, which is commonly evaluated by change in properties after extended heat exposure; it is due rather to the decreasing strength characteristics of the materials at increasing environmental temperatures.

The purpose of this paper is to report the results of somewhat unconventional tests intended to determine the functional properties of silicone rubber seals in aircraft engines. Property studies covered are those of compression set, compression-deflection, and polymer reversion. Correlation obtained between composition and performance is presented.

General Considerations

The merit of silicone rubber for aircraft powerplant seals lies in two major regions — extreme high and extreme low temperatures. The area of our investigation was confined to extreme high temperatures, beyond the useful range of the more conventional synthetic rubbers. Choice of tests was dictated by groove configuration characteristics based on static seal applications for aircraft. The most effective and commonly used design for O-rings comprises a rectangular groove which confines the seal under a diametral pinch. This design embraces two basic configurations: the axial pinch or face seal and the radial pinch seal, both of which are shown in Fig. 1. The assembled stretch of the

* Paper presented at SAE National Aeronautic Meeting, New York, April 1, 1959.

High-Temperature Static Seals

H. E. Todd and J. F. Miazga
Pratt & Whitney Aircraft Division, United Aircraft Corp.

seal is 0-5% for the radial pinch design. In axial pinch configurations the seal is preferably designed such that its OD approximates the outer wall of the groove, and stretch is, therefore, negligible, even under operating conditions. Pinch is controlled by groove depth within the range of 15–25% of the cross-sectional thickness of the seal, and provision is made in the groove width to allow for temperature expansion or fluid swell of the seal without exceeding 100% cavity occupation. One modification of the radial pinch design is a somewhat less effective configuration used for internal straight-thread fluid connections, shown in Fig. 2. Fig. 3 illustrates an entirely different type of configuration used for flexible fittings on rigid pipe connections. The seal, which is of relatively large cross-section, is confined within a cavity defined by the tube, the connecting fitting, a ferrule, and a nut which is tightened against the seal to achieve close to 100% cavity occupation. Assembled stretch on the seal is negligible, but seal is subjected to high compressive stress.

Examination of these seal and cavity configurations reveals that tensile strength and elongation are little utilized by seals in their functional capacity, but rather those properties involved by their confinement and compression. Tests presented in this paper are those which correlated most closely with service experience as it has been affected by increased temperature requirements.

Silicone rubber compounds are relatively simple in composition compared with many organic rubber materials. Their inherent stability is such that antioxidants and other additives used in natural and synthetic rubbers are unnecessary. Each con-

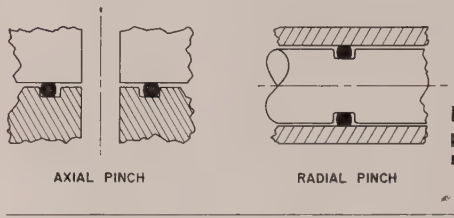


Fig. 1 — Diametral pinch groove configurations for static O-ring seals

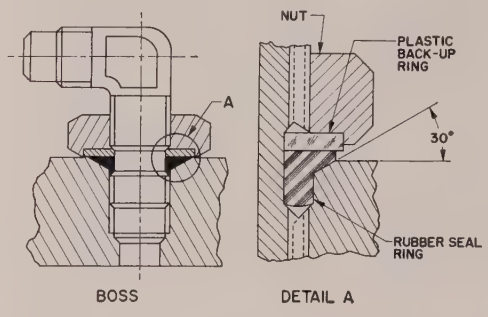


Fig. 2 — Internal straight-thread fluid connection

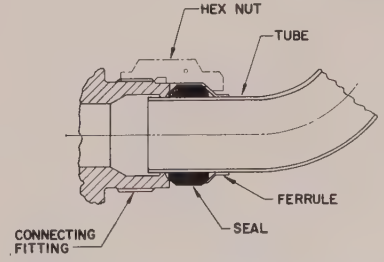


Fig. 3 — Flexible fitting for rigid pipe connection

sists essentially of a silicone polymer or gum, one or more fillers for reinforcement, a catalyst to effect cross-linking or cure, and possibly a processing additive. They are usually mold cured for about 5–15 min at a moderate temperature (240–340 F) but require a higher temperature post cure to stabilize their properties for high-temperature applications. A standard post cure of 24 hr at 480 F was used for all the samples which we have considered in our investigation. Table 1 lists some of the various ingredients used in silicone rubber compounds. The earliest used polymer was a dimethyl silicone which required an additive such as mercurous or cadmium oxide to produce compounds with low compression set properties. Catalysts used with dimethyl gum polymers are benzoyl peroxide and bis (2,4-dichlorobenzoyl) peroxide. The later development of methyl vinyl polymers made it possible to obtain low compression set compounds without the use of additives, and led to the use of less powerful catalyts such as dicumyl peroxide and ditertiary butyl peroxide which in turn effected a closer control of cure. The phenyl containing polymers contribute chiefly to improved low-temperature properties, whereas fluorocarbon and nitrile or cyano silicones accomplish, for the most part, improved resistance to organic fluids. Neither of these properties falls within the scope of this paper.

The earliest used fillers were mineral salts, such as titanium dioxide, ferric oxide, and calcium carbonate, which effected rather low reinforcement or strength. Some improvement was obtained by the use of naturally occurring silica of the diatomaceous earth type such as the Celites and ground quartz such as Neo Novacite. A significant further improvement in reinforcement was attained with commercially manufactured silicas of two general types: those derived from silicates by wet processes, such as Santocel CS ("Aerogel" processed silica) and HiSil X-303 (precipitated silica); and those obtained by the combustion of silicon tetrachloride such as Cab-O-Sil and Dow Corning Silica (fumed silica). This latter type furnishes the highest degree of reinforcement, as demonstrated in Table 2 which lists some approximate average physical properties for typical compounds with various polymers, fillers, and catalysts. Heat-resistance tests for specification testing are commonly conducted on specimens aged 24 or 70 hr at 450 F, at which temperature only slight differences in heat resistance of various compounds are apparent. Our tests were conducted after aging 70 hr at 500 F, both to reveal more clearly the better heat-resistant materials and to approximate more closely extreme service temperatures. It is noted that the original physical properties of "Aerogel" silica and precipitated silica filled compounds are closely similar, but the precipitated silica filled stock exhibits the best heat resistance of the compounds shown. On this basis, precipitated silica was selected as the representative wet process silica for purposes of our tests.

Stacked slab specimens were selected because of easier procurement and closer approximation in thickness to O-ring seals than is the case with compression set discs, which because of their greater thickness do not always exhibit cured properties representative of actual parts. Another reason for the choice was greater accuracy and reproducibility of test results than can be obtained with individual seal specimens. Tests conducted on seals have shown the same approximate trends and relationships, although noticeable differences in absolute values were obtained. Since the scope of the test program was limited to a few representative series of compounds, the relationships presented may not be absolute; however, the data are considered of sufficient cross-section to be reasonably indicative of trends.

Table 1 — Some Compounding Ingredients Used in Silicone Rubber

Polymers	Fillers
Dimethyl silicones	Mineral Type
Methyl vinyl silicones	
Methyl phenyl silicones	Titanium dioxide
Methyl phenyl vinyl silicones	Iron oxide
Fluorocarbon silicones	Zinc oxide
Nitrile or cyano silicones	Calcium carbonate
Catalysts	Natural Silica
Benzoyl peroxide (BPO)	Ground quartz
Bis (2,4-Dichlorobenzoyl) peroxide	Diatomaceous silica
Dicumyl peroxide (DCP)	
Ditertiary butyl peroxide (DTBP)	Manufactured Silica
	Wet process
	"Aerogel" silica
	Precipitated silica
	Combustion process
	(fumed silica)

Experimental Procedure

Specimens for all tests consisted of discs 1.129 in. in diameter (1 sq in. surface area per side) cut from molded slabs approximately 0.075 in. thick and stacked to a thickness of 0.500 ± 0.030 in.

Compression Set — Tests were conducted in accordance with ASTM D 395–55, Method B[1] (constant deflection) with the following modifications:

1. The stacked slab specimen described above was used instead of a standard 0.500 in. thick compression set disc.
2. Specimen was accurately measured for thickness with a dial micrometer and compressed in the test fixture to 75% of the measured free state thickness within ±0.001 in., using shim spacers to provide the necessary adjustment.
3. Specimen was compressed to 75% of its original thickness in all instances, irrespective of its Durometer hardness.

Five temperature-time conditions were used: 22 hr at 350 F, and 70 hr at 350, 400, 450 and 500 F, respectively.

Compression-Deflection — Apparatus consisted of a 0–2000 lb tensile tester equipped with an adapter to provide for compression loading. Attached to the adapter was a loading mechanism comprising a recessed bottom plate to support the specimen and a flat top plate to transmit the load to it (Fig. 4). A dial micrometer was mounted on the instrument so as to measure the travel of the top plate, and hence the applied deflection of the specimen. For elevated temperature tests, a cylindrical heater surrounded the specimen and loading mechanism, the latter being provided with a thermocouple connected to suitable instrumentation for temperature measurement and control.

[1] ASTM Standards, 1958, "Compression Set of Vulcanized Rubber." ASTM D 395-55.

Table 2 — Average Physical Properties of Typical Silicone Rubber Compounds

Silicone Polymer	Dimethyl	Dimethyl	Methyl vinyl	Methyl vinyl	Methyl vinyl	Methyl vinyl
Filler	Mineral	Diatomaceous silica	"Aerogel" and diatomaceous silicas	Precipitated and diatomaceous silicas	Fumed and diatomaceous silicas	Fumed silica
Catalyst	BPO	BPO	BPO	DTBP	DTBP	DTBP
Original Properties						
Hardness, Shore A Durometer	60	70	70	70	70	70
Tensile Strength, psi	450	600	850	800	850	900
Elongation, %	150	70	100	110	125	175
Tear Strength, lb/in.	40	30	45	50	75	100
Heat Resistance, 70 hr at 500 F						
Hardness, Shore A Durometer	[a]	85	76	75	75	75
Tensile Strength, psi	[a]	420	675	690	640	640
Elongation, %	[a]	50	60	90	60	100

[a] No data available.

Specimen was measured accurately for thickness before the test with a dial micrometer, and from this thickness the numerical equivalent in inches was calculated for 5, 10, 15, 20, 25, and 30% compression deflection. The loading mechanism was stabilized at the test temperature within ± 5 F, then specimen was placed in the approximate center of the recessed plate, subjected to an initial load of 3 lb, and temperature conditioned for 1 hr ± 5 min. Room temperature tests were conducted at 70–90 F without a prior conditioning period.

The thickness gage was adjusted to zero and a compressive load was applied to the specimen at a rate of 0.1 in./min until 30% deflection was attained, then immediately released at the same rate to zero loading. This preconditioning cycle was repeated, after which the specimen was subjected once again to a 3-lb load and the dial micrometer readjusted to zero. Compression loading was again repeated, except that instrument was stopped momentarily at the numerical equivalent of each 5% deflection interval to read and record the corresponding applied load. Tests were conducted at room temperature, 350, 400, 450, and 500 F.

Polymer Reversion — Test fixture comprised a small, thick-walled stainless-steel vessel with an ID just sufficiently large (1.130–1.133 in.) to accommodate the stacked-disc specimen (Fig. 5). The height of the vessel chamber was a nominal 0.465 in. slightly less than that of specimen, to provide for light compression loading when cover was assembled. The cover consisted of a flat retaining disc which fitted into a conical seat in the top of the vessel and was secured in place by a screw cap.

Specimen was measured for Shore A Durometer hardness, then placed in vessel. Cover was assembled and screw cap tightened to 25 lb-in. of torque. The assembly was placed in an oven regulated to the desired test temperature and controlled to within ± 5 F, and heated for 6 hr. After removal from oven, assembly was allowed to cool a minimum of 2 hr, then disassembled. Specimen was removed, again measured for hardness and inspected for any pronounced change in condition. Three test temperatures (400, 450 and 500 F) were investigated.

Discussion of Results

Compression Set Properties — Compression set is an inverse measure of the ability of a part to re-

Table 3 — Compression Set at 350 F of Typical 70 Durometer Silicone Rubber Compounds

Compound	Compression Set, % of Original Deflection	
	22 hr at 350 F	70 hr at 350 F
A	105	105
B	25	55
C	15	20
D	12	18
E	19	25

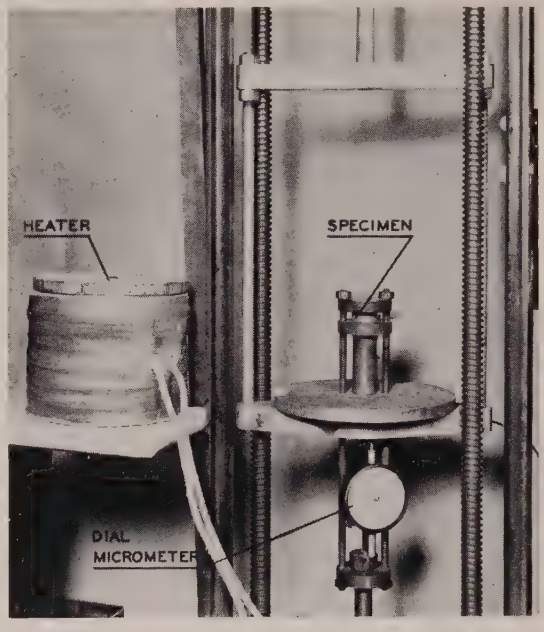


Fig. 4 — Compression-deflection apparatus

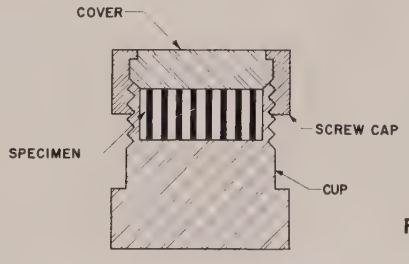


Fig. 5 — Polymer reversion test fixture

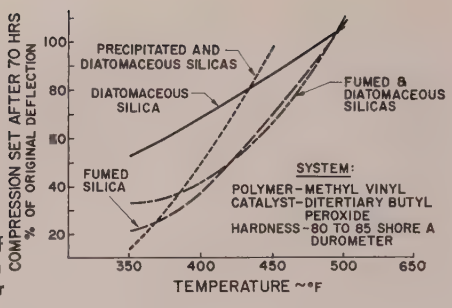


Fig. 6 — Effect of filler on compression set of silicone rubber

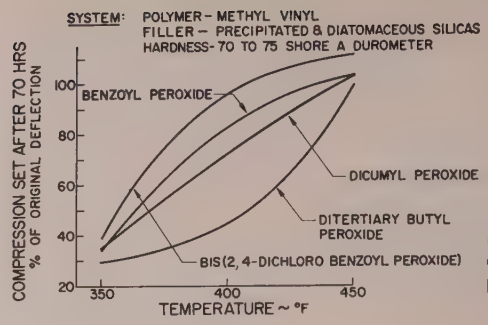

SYSTEM: POLYMER-METHYL VINYL
FILLER — PRECIPITATED & DIATOMACEOUS SILICAS
HARDNESS-70 TO 75 SHORE A DUROMETER
COMPRESSION SET AFTER 70 HRS % OF ORIGINAL DEFLECTION
BENZOYL PEROXIDE
DICUMYL PEROXIDE
DITERTIARY BUTYL PEROXIDE
BIS(2, 4-DICHLORO BENZOYL PEROXIDE)
TEMPERATURE ~ °F

Fig. 8 — Effect of catalyst on compression set of silicone rubber

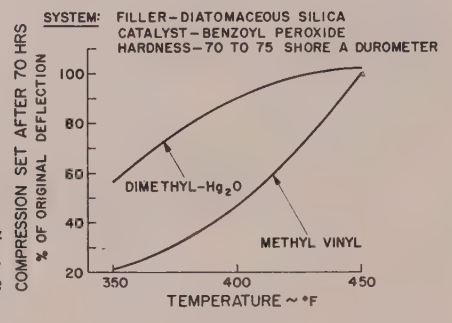


Fig. 7 — Effect of polymer on compression set of silicone rubber

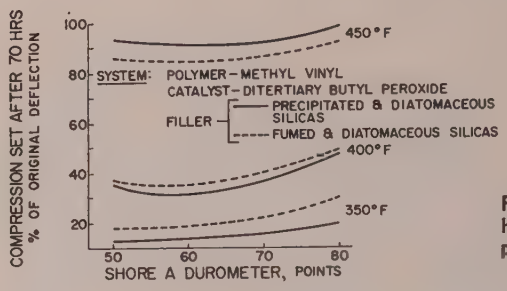


Fig. 9 — Effect of hardness on compression set of silicone rubber

tain effective sealing characteristics, so the lowest possible value for this property is to be desired. Two major environmental factors affect the compression set of elastomeric materials — time and temperature. The effect of time is shown in Table 3 for some representative 70 Durometer silicone rubber componds tested after 22 and 70 hr, respectively, at 350 F. Compound A is typical of the earlier type silicone rubber based on a dimenthyl polymer with an intermediate strength reinforcer such as diatomaceous silica and benzoyl peroxide catalyst. Compound B is of identical composition except for the addition of a low compression set additive, mercurous oxide. The remaining compounds are based on methyl vinyl polymers with mixtures of high- and intermediate-strength reinforcers. Compounds C and D contain precipitated and diatomaceous silicas and are catalyzed with benzoyl peroxide and ditertiary butyl peroxide, respectively. Compound E is similar to D but with fumed silica substituted for precipitated silica as the high-strength reinforcer.

This table shows compression set to be substantially greater after 70 hr than after 22 hr at 350 F for low compression set materials. It also shows the marked improvement in compression set obtained with dimethyl polymer compounds by the use of mercurous oxide and the further improvement afforded by compounds based on methyl vinyl polymers. The seemingly impossible values given for Compound A are not unrealistic for this early type of high compression set silicone rubber tested at 350 F. In addition to taking a complete (100%) set, the specimens underwent heat shrinkage, so that their final thickness dimensions were actually less than those to which they were originally compressed.

Specification tests for compression set are commonly conducted at either 300 or 350 F and at these temperatures lowest compression set was obtained with methyl vinyl base compounds containing precipitated silica as the high-strength reinforcer, as for example compounds C and D in Table 3. At higher temperatures, however, the precipitated silica filled compounds failed to retain their superiority, as shown in Fig. 6 (which demonstrates both the increase in compression set with temperature and effect of fillers). It is apparent from this graph that increased temperature has a severe detrimental effect on compression set. Equally apparent is the more pronounced effect of temperature on compounds containing precipitated silica than on corresponding compounds containing fumed silica. Thus, at temperatures below 375 F the precipitated silica filled stocks have better compression set than do corresponding fumed silica filled compounds, but at 400 F and above, the superiority of the fumed silica compounds becomes increasingly pronounced. The high compression set obtained for compounds containing precipitated silica is believed to be related to their inferior confined heat characteristics which will be discussed elsewhere. Stocks compounded with straight fumed silica as the filler had lower compression set up to about 400 F than did compounds containing a mixture of fumed and diatomaceous silicas; above this temperature the two were essentially equal. The relative temperature insensitivity of compounds containing diatomaceous silica with respect to compression set may contribute to this relationship.

Table 3 has indicated somewhat lower compression set at 350 F for compounds made from methyl vinyl polymers than for those based on dimethyl polymers which contain compression set additives. Fig. 7 shows this same superiority at 400 F, but at 450 F both compounds had similarly high compression set. However, the increasingly widespread use of methyl vinyl polymers is probably due less to

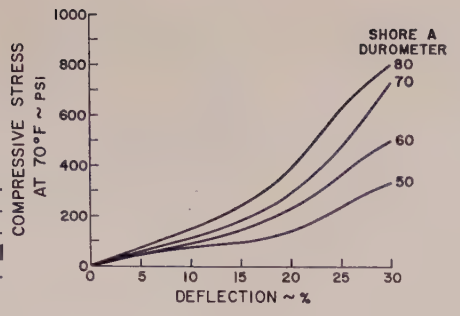


Fig. 10 — Compression-deflection characteristics at room temperature of typical silicone rubber compounds (cartesian)

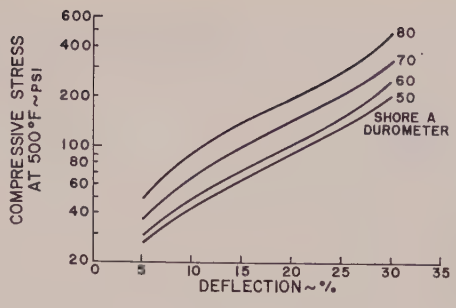


Fig. 12 — Compression-deflection characteristics at 500 F of typical silicone rubber compounds

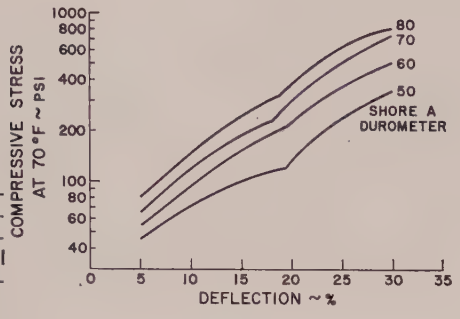


Fig. 11 — Compression-deflection characteristics at room temperature of typical silicone rubber compounds (semi-log)

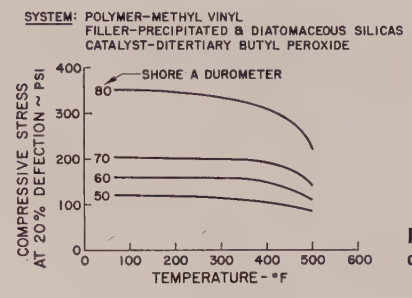


Fig. 13 — Effect of hardness on compressive stress of silicone rubber

their direct contribution to compound properties than to their greater versatility. In addition to eliminating the toxicity hazards attendant with curing of stocks containing compression set additives, methyl vinyl polymers are compatible with a larger variety of catalysts, some of which contribute to improvement of specific compound properties. This is exemplified in Fig. 8, which shows the effect of catalyst on compression set. From the graph it is obvious that ditertiary butyl peroxide (DTBP) is the best of the commonly used catalysts in this respect. Its superiority, though increasingly pronounced between 350 and 400 F, is less noticeable at higher temperatures. DTBP is not suitable for use with dimethyl polymers.

One more factor has been considered — that of hardness. Fig. 9 gives a comparison of the compression set of compounds of varying hardness, prepared with the same gum, filler, and catalyst systems. These curves indicate that lowest compression set is attained in the approximate hardness range of 55–70 Shore A Durometer. They also show again the reversal in superiority of compounds containing precipitated silica and fumed silica with increasing temperatures.

From these curves it appears that lowest compression set at temperatures above 400 F can be expected with a 60–70 Durometer compound based on a methyl vinyl polymer with either fumed silica or a mixture of fumed and diatomaceous silicas as the filler and using a ditertiary butyl peroxide curing system. These data also suggest a temperature limitation of 450–500 F for extended operation of silicone rubber seals, since the compression set of even the best compounds attains or exceeds 100% in this temperature range for 70-hr tests.

Compression-Deflection Properties — Compression-deflection tests measure the stress-strain properties of a material under compression. This test has been more commonly used for evaluating plastics than for elastomers. Tensile stress measurements have been somewhat conventionally employed with elastomeric compounds; however, preliminary tests conducted on silicone rubber stocks for compressive and tensile stress characteristics have indicated no uniform correlation between these two properties, and compression-deflection characteristics were selected for study on the basis that they more nearly simulate one major functional requirement of seals. Tests for this property were confined to methyl vinyl base compounds with various fillers and curing systems, both because they were more readily procured and for reasons of superiority which have already been discussed.

Fig. 10 shows the effect of compressive loading on deflection of some typical silicone rubber compounds at room temperature. Examination of these curves reveals an approach to a straight-line relationship up to approximately 20% deflection, and beyond this point a progressive acceleration of increased deflection with increase in compressive loading. The effect is emphasized by plotting these same curves on semi-log paper, as indicated in Fig. 11 which shows a distinct uniform break in the curves occurring at approximately 20% deflection. This effect is much less pronounced at elevated temperatures because of the greater deflection of the materials at these temperatures, even under low compressive loads. Fig. 12 shows compression-deflection of these same compounds at the extreme temperature of 500 F. It is interesting to note that this 20% transition point coincides closely with the optimum assembly pinch (15–25%) selected for static O-ring seals by aircraft engine designers, based on performance experience. It was also the highest deflection at which an acceptable degree of reproducibility of results was obtained in our tests. For these reasons, the compressive stress at

Table 4 — Polymer Reversion of Silicone Rubber Compounds in Confined Heat Tests

Silicone Polymer	Dimethyl	Methyl vinyl	Methyl vinyl	Methyl vinyl	Methyl vinyl	Methyl vinyl	Methyl vinyl
Filler	Diatomaceous silica	Diatomaceous silica	Precipitated and diatomaceous silicas	Precipitated and diatomaceous silicas	Fumed and diatomaceous silicas	Fumed silica	Fumed silica
Catalyst	BPO	DTBP	BPO	DTBP	DTBP	BPO	DTBP
Original Hardness, Shore A Durometer	70	85	72	73	84	71	81
Loss in Hardness:							
after test at 400 F	0	1	8	3	1	1	0
450 F	0	8	42	30	4	2	2
500 F	3	18	72	70	11	9	4

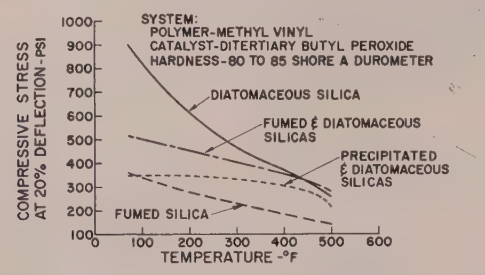


Fig. 14 — Effect of filler on compressive stress of silicone rubber

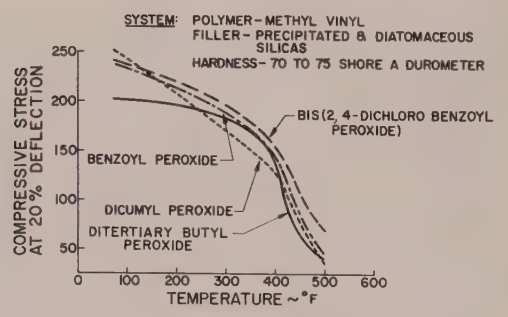


Fig. 15 — Effect of catalyst on compressive stress of silicone rubber

20% deflection was selected as the basis for comparison of various compounds.

Since hardness of silicone rubber is determined by the amount of filler or reinforcer used, it might be expected that compressive stress and hardness would increase somewhat proportionately, provided other factors are constant. Fig. 13 demonstrates that this is indeed the case, although the relationship is not a straight-line function. The four compounds, whose temperature versus compressive-stress curves are shown in this graph, were based on a single formula of 100 parts methyl vinyl gum and 40 parts precipitated silica, catalyzed with ditertiary butyl peroxide, which constituted the 50 Durometer stock. The harder stocks were attained by the addition of increasing amounts of diatomaceous silica. Similar hardness versus compressive-stress relationships were found for other filler systems.

Another relationship shown in Fig. 13 is that of temperature versus compressive stress, which reveals that compressive stress decreases rather slowly with increased temperature to about 350 F, but drops off sharply above this temperature. These curves also indicate that compressive stress is more severely affected by temperature as hardness of the compound is increased. The reason for this phenomenon is explained in Fig. 14, which depicts the effect of reinforcers on compressive stress. This graph shows the extremely rapid loss of compressive stress with increasing temperature for compounds reinforced with diatomaceous silica filler. Since the addition of increasing amounts of this filler was the basis for varying the hardness of the compounds in Fig. 13, it is logical that the harder compounds would exhibit progressively poorer compressive stress versus temperature relationships. The effect is even more pronounced for compounds containing fumed and diatomaceous silicas than for those with precipitated and diatomaceous silicas, which may be due to the rather low contribution made to compressive stress by fumed silica.

Studies to determine the effect of catalysts (Fig. 15) indicate that this variable has relatively little effect on compressive stress. The slightly higher compressive stress obtained for compounds with benzoyl peroxide and bis (2, 4-dichlorobenzoyl) peroxide curing systems is not surprising since these are more powerful oxidizers than ditertiary butyl peroxide or dicumyl peroxide and may effect tighter cures under corresponding time and temperature conditions.

Compressive stress is believed to have a significant relationship to the performance of elastomeric materials for seals. Tipton[2] has concluded on the basis of studies conducted on synthetic rubber compounds that a 20% compressive stress of 400 psi minimum is necessary for satisfactory seal life. Few of the silicone rubber compounds tested have exhibited this high a compressive stress; however, Tipton's tests were conducted at room temperature and were correlated with service life at elevated temperatures. They were also conducted on materials which normally show sharper drops in strength characteristics with increased temperature than do silicone rubber compounds. No quantitative relationship has been established between compressive stress and service life for silicone rubber seals, but there has been striking evidence that high compressive stress contributes to improved service

[2] "Design Data for O-Rings and Similar Elastic Seals," by F. W. Tipton. WADC TR 56-272, July, 1956.

life. This has been especially true for seals of the type used in flexible fittings for rigid pipe connections (Fig. 5) which are subject to high compressive loads.

Confined Heat Properties — There has been observed in many silicone rubber compounds one undesirable property which is not commonly found in most other synthetic elastomers. This property is frequently referred to as polymer reversion. Silicone rubber behaves conventionally when exposed to high temperatures in an unconfined state, gradually hardening and losing strength and elasticity. On the other hand, when it is heated while closely confined, as in a tight container, it tends to soften and become plastic. In extreme cases it attains a putty-like state, somewhat resembling that of an uncured compound.

The mechanism is believed to be one of volatility. Small quantities of volatile breakdown products resulting from heat exposure, if allowed to escape, leave the part relatively unchanged. However, if they are confined in contact with the part, they react with it, splitting polymer chains and producing a deteriorated product. This phenomenon is usually associated with compounds which are not completely neutral, but contain a small excess of either acid or alkali.

The effect of alkalinity is strikingly demonstrated by a comparison of silicone rubber compounds containing precipitated silica and fumed silica fillers (Table 4). Reversion was decidedly more pronounced for stocks with precipitated silica which is derived by a wet chemical process and is prone to retain slight residual alkalinity, not normally present in combustion derived silica. As we have mentioned, the severe reversion of precipitated silica containing compounds is considered a contributing factor to their poor compression set properties above 400 F. Compounds containing "Aerogel" processed silica have been found to have a similar but less pronounced confined heat effect, but the cause in this case is residual acidity. The reversion tendency of methyl vinyl gum compounds filled with diatomaceous silica was also somewhat greater than those with fumed silica, while results were intermediate for compounds containing mixtures of the two fillers. By contrast, dimethyl gum compounds with diatomaceous silica showed remarkably good reversion resistance, which emphasizes the fact that polymer reversion is not necessarily caused by any one ingredient, but is rather a function of the degree of neutrality attained in the finished, cured compound.

Examination of Table 4 reveals that polymer reversion is strictly a high-temperature phenomenon. The data in this table have been limited to the temperature range of 400–500 F; but investigation has shown that below 400 F reversion is almost negligible for any of the compounds, whereas tests above 500 F have resulted in severe reversion of all compounds tested. Obviously, reversion resistance is important for high-temperature silicone rubber seals which are subject to physical confinement.

The mechanism and manner of reversion which have been described may also account for the unsatisfactory behavior of silicone rubber compounds in jet engine oil. Their resistance to this oil is moderately good to approximately 300 F; at higher temperatures they deteriorate in much the same manner as has been observed in confined heat tests. Moreover, this deterioration is not a bulk oil effect, but can occur in the presence of even small quantities of oil vapors, with the severity of deterioration increasing with temperature. Jet engine oil is composed essentially of diesters which are likewise subject to deterioration in this same temperature range, forming, among their breakdown products, dicarboxylic acids. It is believed likely that these acids cause the observed deterioration of silicone rubber materials by upsetting their neutrality in the same manner as do acid or alkaline constituents in the compounds.

Summary and Conclusions

Examination of conventional physical property data for silicone rubber compounds and the results of our somewhat less conventional tests emphasize the difficulty of encompassing all the desirable properties for seal materials in a single compound. Certain generalizations can be made which may serve as a guide in the selection of materials for specific applications.

1. Compounds of 60–70 Durometer hardness provide the lowest compression set, but harder compounds are superior for high compressive stress requirements.

2. Compounds containing precipitated silica as a filler show generally superior resistance to open heat exposure and have moderately high compressive stress characteristics. They are also superior in compression set properties below 400 F; above this temperature compression set increases rapidly, and these compounds are subject to reversion under conditions of confinement.

3. Compounds with fumed silica fillers are somewhat superior in original physical properties, in resistance to confinement at high temperatures and in compression set at temperatures above 400 F. Resistance to heat exposure is intermediate and compressive stress properties are rather low.

4. Diatomaceous silica fillers, when used alone, furnish only moderate reinforcement to silicone rubber compounds. When used in conjunction with higher reinforcing type fillers, they contribute to high compressive stress properties at low and moderately high temperatures, and do not appear to affect compression set properties adversely.

5. Ditertiary butyl peroxide is the best of the catalysts for low compression set and may contribute slightly to improve heat resistance. In other respects none of the catalysts shows distinct advantages, except, perhaps, with respect to processing and curing of compounds.

6. The pronounced effect of heat on nearly all the properties of silicone rubber compounds indicates a maximum temperature limitation of 500 F for their successful use in seal applications.

Acknowledgments

The authors wish to express their appreciation to Dow Corning Corp.; Silicones Products Department, General Electric Co.; Silicones Division, Union Carbide and Carbon Corp.; and Acushnet Process Co. for their cooperation in furnishing background information and for sample compound submissions; also to Dr. D. P. Spalding of General Electric Co. and W. H. King of Acushnet Process Co. for their consultation and advice.

THE CONCEPT of an electric motor mounted inside the rim of a large wheel provides several advantages: flexibility, weight and cost reductions, and adaptability to established operating and maintenance patterns.

A heavy-duty traction motor drive has been designed that eliminates the need for mechanical drive lines, differentials, and hydraulic torque converters. This paper describes the gear train, lubrication system, brakes, and ventilation of such a drive.

Also discussed is the engine-generator set for an electric wheel motor.*

ELECTRIC for

H. J. McLean and H. Vitt
General Electric Co.

WITH INCREASING SIZES of vehicles and payloads, the development of drive systems has tended to lag behind that of the vehicles themselves. Years ago, our company, with its background of experience in electric traction on locomotives, saw in the concept of electrically powered motor wheels a basically simple means of filling this gap, and first applied motorized electric wheels to the "Snow-cruiser" built for Admiral Byrd's third Antarctic expedition in 1939 (Fig. 1).

Over the years, nearly all the large off-highway vehicle manufacturers have discussed with the company the possibility of providing an electric transmission to replace the mechanical or hydraulic drives normally used. However, electric-drive designs using conventional locomotive equipment proved to have excessive weight and high first cost.

On the basis of these inquiries, a detailed design study was begun in 1954 to define the exact requirements of a transmission and to determine the degree to which these requirements could be satisfied with rugged traction equipment designed specifically for application to off-highway vehicles.

The Problem

As vehicles are made larger and larger in an effort to reduce operating costs, maneuverability becomes more and more difficult. Finally, the point is reached where it becomes imperative that the vehicle be of the tractor-trailer type. Getting power to all wheels of such a vehicle presents a difficult mechanical transmission problem. This problem of translating engine horsepower into effective rim pull on all the wheels of this type of vehicle is a natural application for electric drive with its inherent ability to locate the driving motors at the point of use.

To meet this need a heavy-duty traction motor drive has been designed that eliminates the need for mechanical drive lines, differentials, and hydraulic torque converters. Four of these motors can deliver the full output of engines rating up to 1600 hp for traction. The system is designed to provide 60% adhesion on all wheels at standstill and full horsepower utilization over most of the speed range to a maximum of 35 mph. It serves to make electric drive a practical means of economically providing power on every wheel of a single-engine, tractor-type, off-highway vehicle. In fact, power for traction can now be applied to each and every wheel of any off-highway vehicle anywhere a set of flexible cables can be run.

Another method of reducing operating expense is to increase the vehicle payload. Here the weight and space requirements of the drive are important. Materials and space should do double duty wherever possible in order to minimize the tare weight of the vehicle. Electric drive is well suited for this, since the size and shape of the magnet frame required for a traction motor easily conforms to the large, relatively unused space available inside the rim of off-highway tires. Hence, weight and space can be most efficienly utilized by placing the traction motor in this space and using its frame as the axle of the wheel as shown in Fig. 2.

This scheme eliminates the conventional axle, shortens drive lines, and makes a compact driving unit. The motor is also freed of limitations imposed by vehicle design, making it easily adaptable to any large, off-highway vehicle. This configuration also places the items needing most frequent maintenance in the most accessible position—on the outboard end of the wheel rim.

Motorized Wheel

The motorized wheel is basically a heavy steel barrel joined to a large mounting flange. The motor and brake are mounted within this heavy steel bar-

* Paper presented at SAE National Farm, Construction, and Industrial Machinery Meeting, Milwaukee, Sept. 15, 1959.

DRIVE

Off-Highway Vehicles

rel, which doubles as the vehicle wheel axle and the motor magnetic belt, or magnet frame (Fig. 3). A d-c series-wound motor of rugged transportation-type design is used, having the pole pieces for the field structure bolted to the inside of the hollow axle. The armature is carried on a concentric shaft which drives the power gear train through a floating pinion. An airplane-type disc brake is mounted on the commutator end of the shaft.

The wheel bearings are placed around this axle (or magnet frame). They are large-diameter, taper roller bearings, similar to the type used in steel mill rolls under heavy load and shock. The size used can sustain the heavy loads imposed on a motorized wheel of this configuration, and still have capacity for long life.

The mounting flange to which the motor barrel is attached also forms part of the gear-case structure. The motor armature is splined to a short sun pinion shaft whose teeth engage three large gears placed in planetary fashion 120 deg apart. This first stage of reduction is located on the vehicle side of the mounting flange wall in a recess in the vehicle body. Each of these three planet gears is mounted on a pinion stub shaft whose bearings are in the fixed motor mounting flange. These fixed planet pinions emerge through openings in the flange-gear case wall to drive an internal ring gear, thus accomplishing the final reduction. The ring gear, in turn

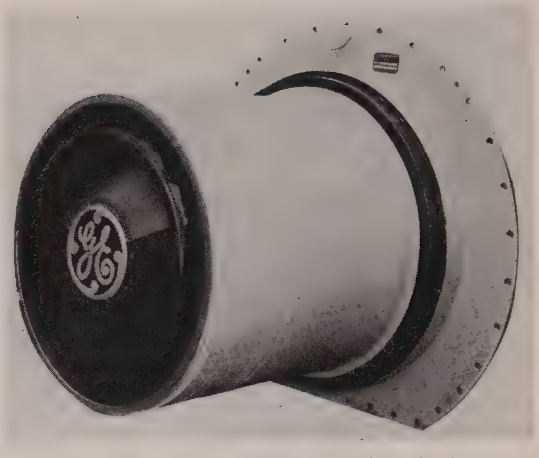


Fig. 2 — Traction motor inserted in wheel

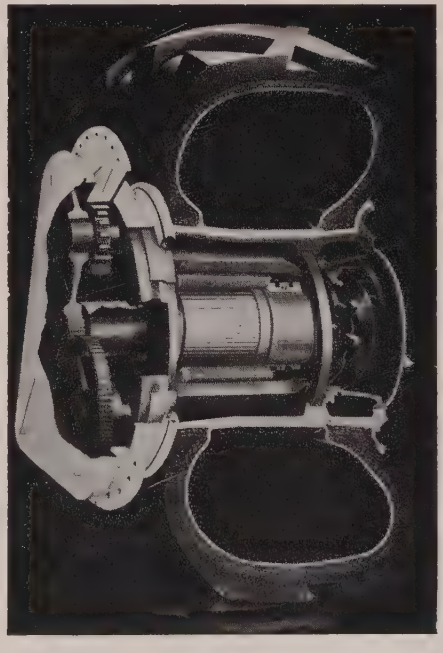


Fig. 3 — Motorized wheel, showing details of motor and gear train

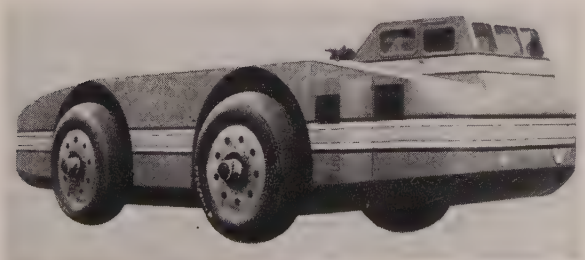


Fig. 1 — Snowcruiser designed for third Byrd Antarctic Expedition, 1939

drives the wheel rim through a spline connection. The mounting flange supports the wheel loading through the magnet frame, provides a means of attaching the motorized wheel to the vehicle, and acts as part of the gear case. The tire, rim, axle, motor, gearing, and brake form one complete unit. In effect, it is possible to bolt four motor units onto any box containing engine, generator, and control equipment and have a truck. In case of major motor or gear failure, the entire motorized wheel unit can be readily replaced in the field and the vehicle kept in operation.

Power Gear Train—When the d-c series-wound traction motor, with its inherent characteristic of high torque at low speeds is coupled with a high gear reduction, tremendous starting torque can be generated. To translate the high torque generated by the motor to that required by the vehicle, a gear reduction of approximately 40/1 is employed. This provides vehicle speeds up to 35 mph and a maximum tractive effort of 32,400 lb, corresponding to 60% adhesion with 54,000 lb on the 44.5–45 tire. With full horsepower utilization over most of the speed range, this transmission provides a highly efficient, smooth flow of power to all wheels.

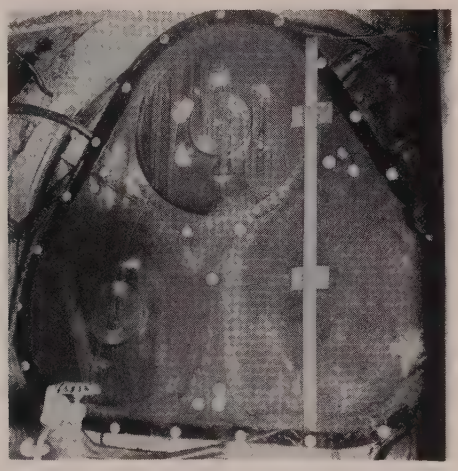


Fig. 4—Splash-lube system

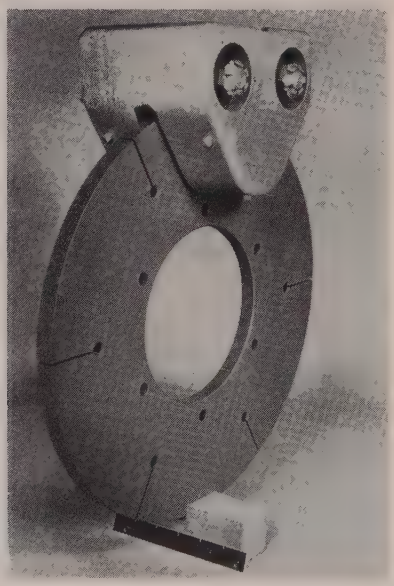


Fig. 5 — Airplane-type disc brake used for parking and emergency service

To transmit these high torques through a single gear of reasonable face width would result in excessive tooth loads, so the obvious solution was to use a modified compound planetary system, with output torque provided by three pinions of normal face width. The planet gears are not allowed to rotate as planets around the sun pinion, but are "fixed in space," so to speak, by building them into the motor frame. The planet carrier is eliminated and the fixed planets drive a movable ring gear. The gearing was placed on the inboard or vehicle end of the motor.

With such high torques to be transmitted it is imperative that each gear tooth carry its share of the load. Unequal load distribution among the gears would result in premature tooth failure and short gear life.

The high-speed gearing is composed of an 18-tooth, 3.85-diametral pitch sun pinion, driving three 87-tooth gears. This tooth combination provides a hunting tooth, thereby improving life and reducing gear noise. Equal load distribution between the gears is accomplished by having the sun-pinion float in its meshes. The three 120-deg spaced meshes support the gear end of the sun pinion shaft. Tooth-separating forces demand the pinion equalize the loads on itself and thus divide the load equally between the gears. At each tooth engagement the pinion runs on a different center, depending on tooth-spacing accuracies, profile error, and the like. This center is dictated by the resultant of the three force vectors from the gear meshes. The floating action of the sun-pinion shaft is accomplished by a sliding-fit splined connection between it and the motor armature.

In the low-speed or final reduction the tooth loads are higher, and hence it is the more critical of the two meshes. The 15-tooth 3.2-diametral pitch pinions mesh with a 111-tooth internal gear. Here the problem is a little more difficult. The ring gear must be designed to: (1) float sufficiently to equalize the loads of the three planet pinions; (2) engage the wheel rim positively to transmit driving torque; and (3) not transmit wheel or axle loads through the planetary gear meshes and thus overload both gearing and gear bearings.

The splined drive ring attached to the rim is designed to help fulfill these requirements. Its cylindrical shell has a section flexible enough to allow distortion by tooth-separating forces and to prevent transmission of heavy axle loads to the gearing. Thus, the ring gear floats by distortion, both of itself and the splined drive ring, to equalize tooth loads.

This built-in flexibility makes the gear train self-adjusting to the gear loads, resulting in improved performance and increased gear life.

Lubrication—The gearing and bearings are splash-lubricated by oil, except for the motor armature and outer wheel bearings, which are grease-lubricated. Essentially, the splash-lube system operates by lifting oil to the top of the gear case by means of the two lower planet gears. On its return path the oil is directed to appropriate locations by means of channels and gutters. Two gutters feed oil to the ring gear which carries it to the top of the low-speed gear case for a second cycle of lubrication.

Oil action in the high-speed gear case of the motorized wheel used to design the splash lube system is shown in Fig. 4. The oil reservoir, channels,

and feed lines are of ample size to assure copious flow and prevent clogging.

An oil sump located beneath the two lower planet gears also serves to collect wear particles and heavy sediment that becomes mixed with the oil. This sump is covered by a plate with nozzle-like openings pointing downward, thus forming a "screen" that creates a dead oil area and serves to keep particles which may damage bearings or gearing out of the oil circulation.

Brakes—The motorized wheel has two brake systems. The primary system is the electric or dynamic brake. The secondary system consists of a mechanical airplane-type disc brake on each motor.

Normal service braking is performed by the electric brake in a manner that will be described later. This system is capable of developing a braking effort of approximately three times the available traction horsepower, and has no wearing parts, such as linings or drums that require periodic replacement.

The secondary system is used as a parking brake and as an emergency brake in case of engine failure. While it may also be used simultaneously with the dynamic brakes to develop an unusually high retarding force, it is not intended for use as a day-to-day service brake.

The secondary brake is composed of two main elements as shown in Fig. 5—a flat slotted disc splined to the motor armature shaft, and a double-acting hydraulic cylinder assembly rigidly bolted to the motor frame. The two pistons are headed by a brake lining material which in operation act in a C-clamp action on opposing sides of the brake disc. The brake cylinders are hydraulically operated, either directly or from an air-over-hydraulic power cluster. Since this brake is used for parking, it may be desirable to provide an accumulator to maintain brake cylinder pressure in the event of slow air system leakage. This would be particularly desirable if the vehicle were to be parked overnight on a steep grade.

The brake operates very smoothly since it is not self-energizing. The braking action is transferred to the wheel with no danger of damage to the gears.

This disc-type friction brake has an advantage over the conventional drum and shoe brake in that there is no brake "fade." Since the linings bear against the sides of the disc, the expansion caused by heating will not affect the contact area as in a drum brake. The disc is of a special alloy, designed to operate under normal conditions at about 600 F with a maximum operating temperature of about 1000 F. Maximum kinetic energy capacity is 2,890,000 ft-lb per brake for any one stop. To raise the disc to the normal operating temperature requires 1,735,000 ft-lb; more than enough energy to bring a 4-motor, 100-ton vehicle to a stop from 30 mph. Torque developed by the brake is 20,000 in.-lb at 590-psi hydraulic pressure.

The brake disc is situated in the motor cooling airstream, thus helping to keep operating temperatures low. Air from the motor impinges directly upon the inner side of the disc. The outer side is cooled by a radial vane-type fan cast integral with the brake disc mounting hub. This cooling will enable the operator to reapply the brakes much sooner than if no artificial cooling were used.

The brake cylinder and disc are easily serviced by removing the motor hub cap. No part is too heavy for one man to handle easily. Brake shoes need only be inspected during the regular monthly motor brush inspection.

Ventilation—The concentration of so much power in a relatively small space and weight is made possible largely through the use of high-temperature insulation and generous ventilation, in accordance with modern traction equipment design practice.

Because of the location of the gear case, introducing air to this motor was a major design problem. The method used was to blow air into the motor by going under the gear case, mounting the motor in a plenum chamber built into the vehicle.

This is accomplished by welding the motor barrel to a mounting flange and attaching the gear case to the rear or inboard side of the flange, raising the gear case up free of the flange by means of ribs or stilts, as shown in Fig. 6. The motor is then bolted to a cavity on the vehicle, to which air is introduced at another point. The air flows over the gear case, around the sides, under the gear case, and through the motor field and armature, over the brake disc, and out through the hub cap.

This design has the following advantages:

1. No external ducting need be connected to the motor, all air connections being made automatically when the motor is bolted to the vehicle.
2. Cooling of the gear case itself by airflow over the entire back and about half of the front, thus carrying away the heat buildup in the high-speed gearing and cooling the lubricating oil reservoir located in the gear case.
3. No ducts run through the gear case, eliminating a potential source of oil leakage and keeping the oil capacity of the gear case at a maximum for heat dissipation purposes.

The air outlet is spray or splash-proof and with a 44.5–45 tire allows the vehicle to negotiate water 52 in. deep without danger of flooding the motor. Application of watertight covers to the air outlets would permit the vehicle to operate for short periods of time with the motors completely submerged. For continued submerged operation a "snorkel" arrangement could be developed to keep the motor from overheating.

Some Interesting Design Points—In designing a machine of this size and complexity many problems —some obvious and some almost obscure—must be solved. All parts must work toward providing a

Fig. 6 — Rear view of motor showing air passages between gear case and mounting flange

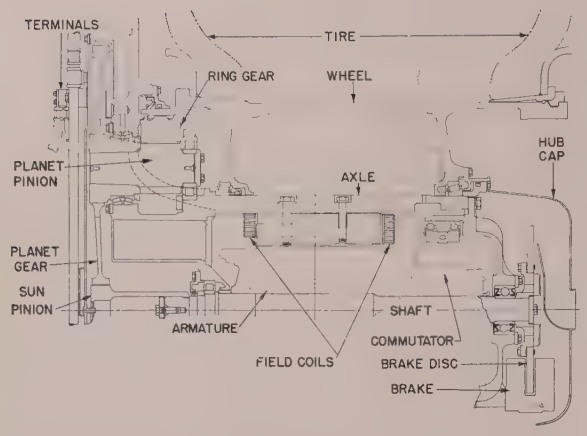


Fig. 7 — Simplified longitudinal section of motorized wheel

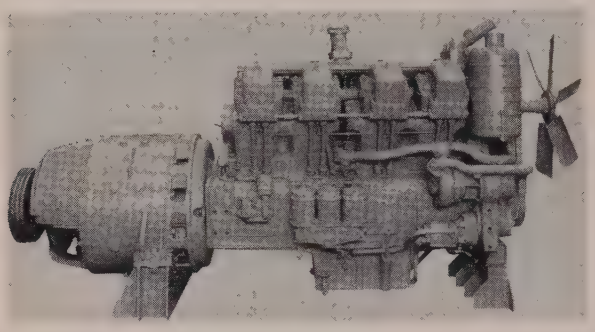


Fig. 8 — Representative engine-generator set

smooth flow of power from the electricity fed into the motor to the traction at the tire tread. All components (Fig. 7) must be so related that not only will the main drive line function properly, but also all its supporting and aiding parts will contribute toward the main task of providing smooth, reliable tractive effort.

Examples of a few of these related components and their functions will serve to illustrate their importance.

In this design of wheel, oil seals of larger diameter than usual are employed. The gear case oil seal, with a sealing diameter of 42½ in., statically must seal against a head of approximately 8 in. of oil, and dynamically against 4 in. of oil. The inboard wheel-bearing seal, although sealing against oil, is bathed only by the oil fed to the wheel bearing, but is required to have the same sealing capacity as the gear case seal.

During the design of the production model of this wheel, considerable time and effort were expended in investigating seal configurations and effects of variables on seal performance. As a result a seal housing design has been developed which gives satisfactory performance under service conditions. Sealing is effected by a double seal arrangement. One rubbing seal to keep oil in, and another in series with it to keep dirt out.

Another item requiring consideration is the effect of the expansion and contraction of the air under the rim where it is subject to wide temperature variations. With a 45-in. rim there is a volume of about 14 cu ft of air in this space. When the vehicle is shut down or left idle for a period of time, this air space will approximate ambient temperature. During operation, however, it is heated by the motor frame, the hot oil in the gear case, and the hot air in the tire. Since the volume enclosed remains constant (Fig. 7), a rise of temperature would tend to develop pressure in the enclosed space. When the motor cools off the process will be reversed, and a respiratory action would tend to result.

This condition is relieved by a small hole drilled through the motor frame which serves as a pressure equalizer.

Location axially of the sun-pinion shaft was another design problem. Since this shaft must be free to move radially in any direction it chooses at its gear end, it cannot be rigidly connected to the armature shaft.

To restrict axial movement in one direction, the spline teeth of the sun-pinion shaft impinge against a shoulder of the armature shaft. To hold it axially in the other direction, an alloy steel plate is bolted to the end of the armature shaft to provide a thrust surface for the spline teeth. The end of the armature shaft is relieved to support the plate near its outer edge, thus providing ample follow-up to maintain the bolt tightness. Since the retaining members are either attached to, or a part of the armature shaft, there is no relative motion between the sun-pinion shaft and the retaining parts.

This design has a distinct advantage over absorbing the thrust by a thrust bearing or washer against which the shaft must rub at considerable velocity. It also allows assembly and disassembly of the sun pinion shaft by simply removing one bolt.

By making the shaft hollow to allow use of the plate retention device (Fig. 7), an easy, effective means was provided for lubrication of the spline. Oil is introduced into the hollow pinion and centrifugally forced through the spline teeth.

Ratings—A traction motor is not conventionally rated by horsepower like most engines and other prime movers. Several designations are used when stating the rating of an electric traction motor. Perhaps the most important of these is the so-called "continuous rating." This means the maximum current which a motor can carry continuously at rated voltage, with full field excitation and specified ventilation without exceeding the allowable temperature rise. Translated into terms that may be more familiar, it is the highest tractive effort the motor can develop continuously without overheating.

Sometimes, as with this motorized wheel, an electric traction drive is also rated in "per cent adhesion." If we say the motorized wheel will develop 60% adhesion at standstill with a loading of 54,000 lb per wheel, we mean that the maximum allowable motor current at standstill will produce a torque that will give a tractive effort of 32,400 lb per wheel. At its continuous rated current this motor will produce a certain torque which will give a tractive effort of 10,800 lb per wheel. This tractive effort divided by the rated weight per wheel corresponds to 20% adhesion. The motor is then said to be rated 20% adhesion continuously. A vehicle having the traction motor rating at a per cent adhesion high enough to be near the slipping point of the wheel is said to be "motored up to adhesion."

The motor also has shorter time ratings at which it may run. These will be illustrated in the speed tractive effort curves presented later. These shorter time ratings are the high torque outputs used to start and accelerate the vehicle.

Finally, a word about how horsepower ties in with these ratings and motor performance. The tractive effort developed by a d-c series-wound motor varies with the input current. The speed on the other hand, varies with the voltage. In any particular instance a certain amount of current and voltage (the product of which represents horsepower) is required to keep the vehicle moving at a given speed. For example, let us suppose a vehicle moving up a 5% grade requires a current of 500 amp and a voltage of 300 volts on the motors. The motors are developing sufficient tractive effort, and since current determines tractive effort—the effect of increasing horsepower to the wheel woud be to increase the voltage only, thus speeding up the vehicle. The only increase in tractive effort (or current) required would be that needed to overcome any increased rolling resistance at the higher speed.

Electric Drive System

Electric power for the wheel motors is provided by an engine-generator set. The generator is usually designed to mount directly on the engine as shown in Fig. 8. An adapter is used to assemble the generator frame to the engine, and the armature is driven by the engine flywheel through a disc coupling that is stiff radially and torsionally, but soft axially.

Powerplant Design Considerations—A torsional study of the crank shaft-armature assembly is always required on new applications to determine if any critical vibration frequencies occur in the operating speed range. The close-coupled, single-bearing generator design helps to keep the resonant frequencies high but, if necessary, changes in the coupling, engine firing order, or related components can be made which may shift the frequencies at which resonance occurs beyond the operating range.

Another important consideration in applying and installing traction generators is an adequate supply of cooling air. As with the motor, forced ventilation permits a substantial increase in the current and power rating of the generator and proves to be an economical way to increase its capacity. The cooling air is generally supplied by an externally driven fan and is cleaned to reduce the amount of dirt carried into the machine.

The purpose of ventilation is to remove the heat produced inside the generator as the result of mechanical, electrical, and magnetic losses. The electrical designer must take into consideration the effect of these losses on equipment first cost, life, and maintenance in developing a generator design. He must also produce a machine that will meet the current and voltage requirements of the traction motors without exceeding the speed and torque limitations of the engine that will drive it. The current, speed, and voltage requirements of the motor load establish many of the generator electrical design parameters. For example, the continuous current rating of the generator should be determined by that of the motor combination it is to drive. Since motor speed is, to some extent, dependent upon generator voltage the maximum generator voltage must be high enough to permit the motors to meet the speed requirements of the application.

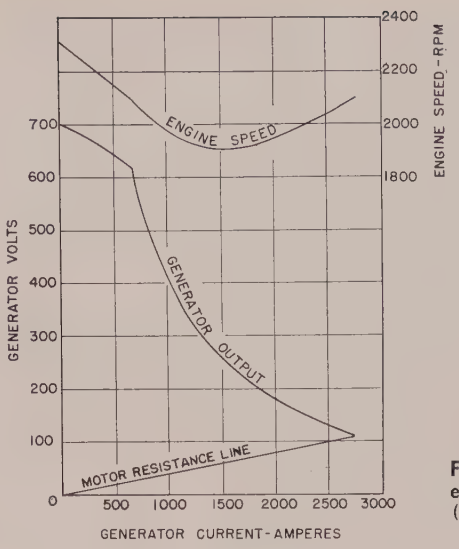


Fig. 9 — Engine-generator characteristics (600-hp, 2100-rpm engine)

The remaining generator design parameters are established by the engine. For instance, the rated speed of the engine is important in determining the generator field requirements. The terminal voltage of any d-c machine can be expressed approximately by the relationship:

$$V \approx K_1 N I_f \qquad (1)$$

where:

V = Terminal voltage
K_1 = Constant for the machine under consideration
I_f = Field current
N = Armature speed

This equation states that voltage is proportional to the product of armature speed and field current. From this it follows that an increase in engine speed will make it possible to get the same generator terminal voltage with lower field current. If, on the other hand, Eq. 1 is applied to a motor, the armature speed can be determined for any terminal voltage and field current.

To obtain satisfactory operation over the wide speed range required in vehicle applications, it is necessary to change both motor voltage and field current. The obvious way to change motor voltage is to change the terminal voltage of the generator. However, there is one other requirement that must be met when generator voltage is changed.

This requirement is most easily understood by examining a very simple equation:

$$P = VI \qquad (2)$$

where:

P = Power, watts
V = Terminal voltage
I = Current, amp

Since the engine driving the generator develops a constant power output for a fixed governor fuel setting, it is desirable that the VI product of the generator remain practically constant. Therefore, as the generator voltage increases, its output current must decrease. If this does not occur the engine will be overloaded and its speed will drop. On the

other hand, if the generator output current decreases too rapidly as its voltage increases, the engine will be unloaded. Its speed will increase, and it will run "on the governor."

By designing the generator excitation system so that the generator power demand decreases more rapidly than the engine output as speed drops, a good match can be obtained between generator demand and engine output. Terminal characteristics for a typical generator driven by an engine rated 600 hp at 2100 rpm is shown in Fig. 9.

Motor Output Characteristic—The power output of the generator is only useful when it becomes speed and torque at the wheels of a vehicle. As already pointed out, the electrical factors that determine motor speed are terminal voltage and field current. The relationship between the electrical factors that determine motor torque is:

$$T \approx K_2 I_f I_a \qquad (3)$$

where:

K_2 = Constant for the machine under consideration
I_f = Field Current
I_a = Armature current

This equation takes a special form for motors generally used in traction applications since they are of the series type. They are so called because the motor field is connected in series with the armature and carries the same current. This is easily seen from the elementary sketch of Fig. 10. For this special case, Eq. 3 can be modified to read:

$$T \approx K_2 I^2 \qquad (4)$$

where:

I = Motor current

If Eq. 1 is rewritten for the motor:

$$V = K_1 N I$$

and both Eqs. 1 and 4 are substituted into the power Eq. 2, the result is:

$$P = \frac{K_1}{K_2} NT \qquad (5)$$

where:

P = Generator output power
K_1 and K_2 = Motor constants
N = Motor speed
T = Motor torque

It can be seen from Eq. 5 that, theoretically, a series motor supplied with power by a generator with the proper output characteristic will trade speed for torque to maintain constant power at the motor shaft.

While theoretically perfect motors and generators are not economically feasible, practical machines can be built that have very useful output characteristics. For example, Fig. 11 shows the speed torque curve for a set of four series motors designed for vehicle application when they are connected to a generator having characteristics as shown in Fig. 9.

This combination, installed on a vehicle with approximately 40/1 gearing and tires of 103-in. rolling diameter, will develop the tractive effort curve of Fig. 12. Tractive effort is used here as the total pull developed at the wheels. It is determined by deducting gear losses from motor output and referring the result to the tire diameter. Drawbar pull can be calculated by deducting from tractive effort the power required to overcome rolling resistance and other external losses.

Power transmitted equals the tractive effort developed at a wheel tread multiplied by the speed at which it is turning. It is desirable, of course, to transmit full power over the widest possible range of speed. If necessary, the speed range over which a series motor will transmit full power to the wheels, can be extended by connecting a resistor around the motor field. This produces an effect similar to that obtained by shifting gears with a mechanical transmission. The resistor weakens the field by diverting some of the armature current around it. Since Eq. 1, $V = K_1 N I_f$ must still be satisfied, the motor must run at a higher speed. In practice the resistor connection is made automatically, and the operator

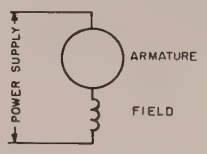


Fig. 10 — Series motor connections

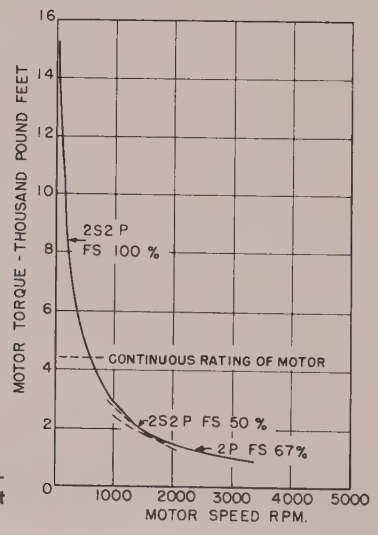


Fig. 11 — Speed-torque characteristics of 4-motor equipment with 600-hp engine

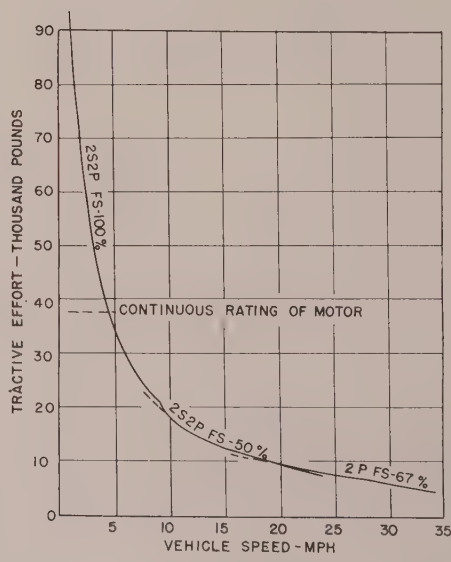


Fig. 12 — Vehicle speed-tractive effort curve 4-motor equipment with 600-hp engine

can barely notice the transition.

Motor Arrangements — With an electric transmission power can be supplied by means of flexible electric cable to as many wheels as necessary, even though some of them are remote from the power source. In this respect, the electric transmission is unique, since other types of transmissions are inherently limited to two powered wheels. The problem of supplying full power to several motors leads to a consideration of the several possible ways in which they can be connected.

Turning again to the example of a vehicle with four motors, we find that there are three basic ways in which these motors can be connected. One way would be to connect them all in series across the generator as shown in Fig. 13.

In this connection, the voltage across any particular motor is one-fourth of the generator voltage. For the type of machines we are considering, this connection would require extremely high generator voltage to drive the motors fast enough to obtain satisfactory vehicle speeds. This introduces a number of difficult problems in the generator design that are expensive to solve.

Another disadvantage of the series connection stems from the fact that the speed of a series-type motor increases as its load decreases. Therefore, if one wheel should slip, its motor would speed up, and tend to cut off the current to the other motors. As a result, power to all wheels would be reduced and the vehicle might stall. In effect, this is similar to what might happen if all wheels were powered from one 4-wheel differential.

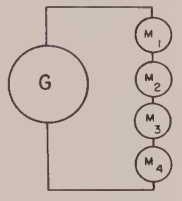


Fig. 13 — Series connection of motors

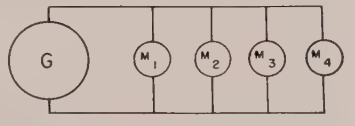


Fig. 14 — Parallel connection of motors

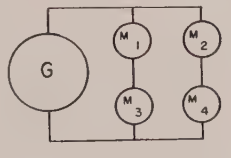


Fig. 15 — Series-parallel connection of motors

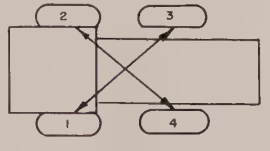


Fig. 16 — Cross-connection of motors on vehicle

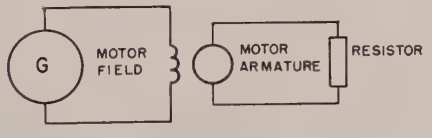


Fig. 17 — Basic connection for dynamic braking

Another way to connect the motors would be to put all four in parallel across the generator (Fig. 14). Now each motor gets full generator voltage, but the generator must supply four times the motor current. While there is no problem about vehicle speed, the generator must have an extremely large capacity to supply the heavy current required to start a loaded vehicle. An advantage of this connection is that any motor that slips only shifts its share of the generator power to the other three motors which continue to supply full torque to their wheels.

The third motor connection scheme is a compromise between the two just described. Two groups of two motors in series are connected in parallel (Fig. 15).

This connection has some of the advantages and disadvantages of both the others. Generator voltage required is half that necessary in the series connection to drive the motors to the same speed. Generator current is half that required in the parallel connection to get the same tractive effort. With this connection, if one wheel slips it affects only the other motor in series with it. The torque of this motor is reduced but the remaining two motors in the other series group can still continue to supply full torque to their own wheels. By cross-connecting the motors from front to rear of the vehicle as shown in Fig. 16, there is little possibility that the vehicle will stall as a result of slipping. Also, motor torques when making a turn with the vehicle are very nearly equal. Experience to date with this connection in actual service has been very favorable.

Dynamic Braking — Although the importance of getting all the available power to the wheels cannot be over emphasized, the job of getting it back out while braking a heavily loaded vehicle is equally important. In handling heavy loads on steep grades, ample braking capacity not only promotes safety, but also reduces cycle time.

The d-c wheel motors offer a means of obtaining braking effort in quantities large enough to affect operating methods materially. If the correct connections are made, a d-c machine can be made to operate as either a motor or generator depending upon the direction of power flow. When electrical power is fed into the machine and mechanical power is delivered at the output, it is said to be motoring. When mechanical power is fed into it and electrical power is delivered at the output, it is said to be operating as a generator. The voltage and torque equations stated earlier apply in both cases.

By making the proper connections to the electric motors at the wheels, they can be made to operate as generators and act as brakes. This requires two changes from the normal motoring connections. First, the motor fields must be reconnected so that generator current runs through them only and not through both the motor armature and field. Second, a suitable means of absorbing electrical power must be connected across the motor armature. The result is a new set of connections that may be represented by the diagram in Fig. 17.

Now power developed by the motor acting as a generator is dissipated as heat in the resistor and the force required to drive the motor retards the wheel.

As usual there are certain limits on such quantities

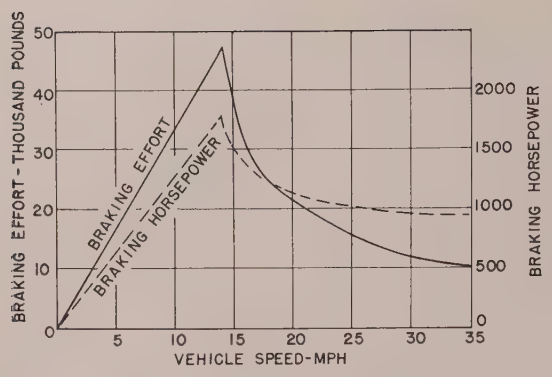


Fig. 18 — Dynamic braking characteristics (4-motor equipment)

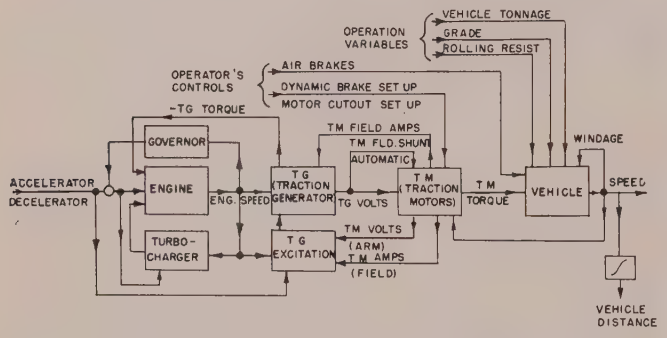


Fig. 19 — Simplified block diagram for analog computer simulation of diesel-electric vehicle

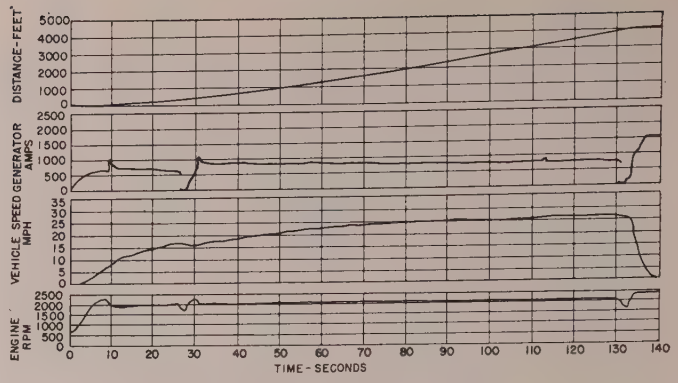


Fig. 20 — Typical computer simulation of performance cycle for 4-motor, 600-hp vehicle of 110-ton gross load

as voltage, current, speed, and torque that must be observed. However, within these limitations it is possible to obtain a substantial amount of braking power, as can be seen from Fig. 18. This represents the performance in braking, of the same 4-motor equipment considered previously in traction.

One important limitation of electrical braking is that it cannot serve as a holding brake since, as shown in Fig. 18, the braking effort vanishes at zero speed. Hence, a mechanical brake must be provided to hold the vehicle at standstill.

Operator's Controls — Disregarding the steering function, which is not included as part of this equipment, the three functions that an operator must be able to control are:

1. Traction power.
2. Braking power.
3. Direction (forward, neutral, reverse).

Traction power is controlled in the conventional manner by controlling the fuel supply to the engine. An accelerator pedal worked by the operator sets the position of the engine governor control arm just as it would on a mechanical transmission. The only additional requirement of the electric drive is that a signal be provided to tell the generator that engine fuel is being increased. This is easily obtained by actuating a switch from the throttle linkage. Once the switch has been operated, all further increase in power takes place automatically.

Braking is generally controlled by another pedal like the accelerator pedal, also worked by the operator. The brake pedal, however, must operate three different devices: (1) like the accelerator, it operates a control switch to set up the proper electric circuits; (2) it drives the operating arm of a variable resistor that increases the amount of braking as the pedal travel increases; and (3) it is connected to the engine governor throttle arm to increase engine speed, as pedal travel increases.

The need for the third function should be explained. The blower for cooling the tractor motors is driven directly by the engine. To obtain maximum braking effort from the tractor motors, it is necessary to run the engine at full speed to get maximum cooling air to the motors. The load on the engine during this operation is light and the engine will be running at high, light-load speed "on the governor."

Analysis of Vehicle Performance — Curves, such as the speed-tractive effort and speed-braking effort characteristics shown in Figs. 12 and 18 are useful in studying the performance of a vehicle when operating on long cycles. Frequently, however, the operating cycle is short enough to make the time spent in accelerating from idle to full power a substantial part of the total. In such cases a transient analysis is required to evaluate the performance of the transmission system. The analog computer proves very valuable for this type of work. It can quickly perform all the necessary mathematical operations and when required, can generate non-linear functions such as engine torque-speed curves.

By using the computer, it is possible to simulate all parts of a system that enter into a consideration of its performance. Sufficient information is generally available on such system components as engines, turbochargers, generators, and motors to make satisfactory simulation easy. Other factors influencing vehicle performance, such as rolling resistance and wind resistance, can only be estimated. The limitation here lies in the validity of the information available concerning these factors and not in the ability of the computer to provide a satisfactory simulation. Generally, the most practical solution is obtained by using the best available estimate for these difficult factors in order to include their effect in the simulation.

When complete, the simulation takes into account all elements that affect vehicle operation including operator's controls, engine governor, turbocharger, traction generator, traction motor, vehicle load, grade resistance, rolling resistance, and windage. Simulation of the operator's controls is made easy by the fact that there are only three. The block dia-

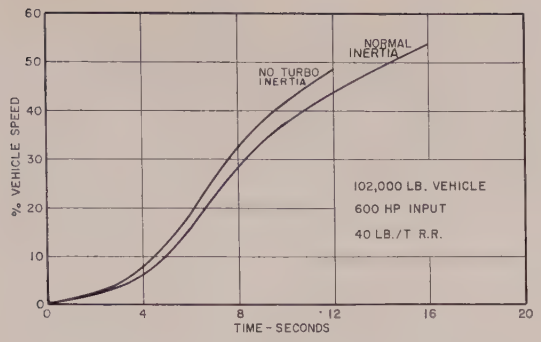


Fig. 21 — Effect of turbocharger time constant on vehicle performance

gram for such vehicle simulation is shown in Fig. 19.

The output of many analog computers is a strip chart run at constant speed on which the displacement of a pen indicates the variation of the quantities being studied. The result is a set of recordings that looks much like oscillograph charts. A typical section of such a chart from the motorized wheel studies is shown in Fig. 20. Here the distance traveled, generator current, vehicle speed, and engine speed are shown for a vehicle with a 600-hp engine and 110-ton gross load running on level ground with 40 lb per ton rolling resistance. Since the speed at which the strip chart runs is carefully controlled, performance as a function of time can be determined easily.

Once the computer setup has been made, it is possible to simulate a change in any part of the system and evaluate its effect. For example, the turbocharger time constant can be changed on the computer and the effect of such a change checked rather easily. By comparing the strip charts, the effect of variation in the turbocharger time constant is readily seen. A cross-plot of this information, as shown in Fig. 21, can be made to display the effect of the change more clearly if necessary.

Other components of the vehicle drive system can be studied just as easily on the analog computer. Thus, this valuable tool permits rapid and economical evaluation of the effect of changes in various components on vehicle performance.

Summary

The concept of an electric motor mounted inside the rim of a large wheel provides new standards of flexibility and performance. It also offers weight and cost reductions over more conventional electric drive schemes that make the advantages of electric power application to a number of wheels and increased braking capacity economically feasible in large off-highway vehicles. Moreover, the motorized wheel vehicle will easily fit into established operating and maintenance patterns. The functions to be controlled by the operator of a vehicle with electric transmission are by no means unique. The devices for operator control can be arranged to require only slight physical effort and practically no need for re-training when switching from a vehicle with mechanical transmission.

Field service tests indicate that the equipment is sturdy and reliable, and that when properly applied it is capable of outstanding performance.

DISCUSSION

Electric Drives Offer Some Problems

— John Hyler
LeTourneau-Westinghouse Co.

SINCE HAVING seen some of the supposedly new inventions that were actually conceived and built centuries ago, it is difficult to say just when the first electric drive for off-highway vehicles appeared. However, we do know that in 1923, our predecessor company designed and built an all-electric drive scraper which was the beginning of large capacity, one-man operated earthmoving machines. At least by 1925, electrically driven shovels were beginning to make their appearance, employing the electric drive for the crawlers of the carrier as well as the upper works.

Our company has produced equipment that will take advantage of the relatively low cost of electric power in some areas by employing a trolley system. This permits our a-c generator to function as a motor for the greater part of the hauling cycle, but allows the diesel engine to take over for any operations away from the trolley. More recently, we have seen the Kenworth truck modified to electric drive by using a d-c two-trolley system for its off-highway hauling operations.

As early as 1950, R. G. LeTourneau, Inc. began to develop and build the d-c electric wheel which is rather similar to that described in the paper. It is used today in their 100-cu-yd earthmovers, 150-ton tree crushers and in their over-land trains. The latter operate on 10 ft diameter tires across desert or arctic areas.

Army Ordnance has also pursued electric drives for some of their vehicles. These have been diesel-electric units driving an electric motor which in turn is connected to a transmission and controlled differential arrangement. Several experimental tanks have incorporated this drive. The Army Engineers have used an electric assist drive on the front wheels of their all-purpose tractor known as the BAT, while retaining mechanical means to propel the rear wheels.

My company has provided electric power on some of their machines for the trailing axle while retaining the mechanical drive for the front axle. The most recent application of this arrangement is on the GOER vehicles which have been developed under the direction of Army Ordnance.

While the components described in this paper do provide an extremely flexible arrangement, we believe that the case for the electric transmission is overstated when the authors refer to all other types of transmission being inherently limited to two powered wheels. The hydrostatic transmission, on which so many people are concentrating, represents just one of several other types or combinations that also permit a fair amount of freedom in vehicle design.

Reference is made to the extensive work done to obtain a satisfactory seal at the 42½ in. diameter location. This would appear to be a likely source for continuing trouble for some off-highway applications.

The safety of the parking brake system depends upon

maintaining sufficient hydraulic pressure for as long as the machine may be parked.

Referring to Fig. 18 of the paper, we notice that the electrical braking effort and braking horsepower both drop off suddenly above 14 mph. Since it is stated that the motors are capable of handling up to 1600 hp, it would seem desirable to try to eliminate this reduction of braking capacity at higher speeds so that the operator will not lose his confidence and unnecessarily limit performance.

I do not agree that vehicles must be tractor-trailer type to be more maneuverable. Our two-axle trucks, for instance, with automotive-type steering and hauling well above their rated 32 tons, can turn in a smaller circle than the normal passenger car. However, the two-axle units with wagon steer, such as produced by practically every manufacturer of earthmoving machines, represents in my opinion the most maneuverable unit to date.

We agree with the statement that a method of reducing operating expense (per unit carried) is to increase the vehicle payload. This may be approached in two ways:

1. Decrease the weight of the empty machine so that the same size components can deliver more payload. Although the system presented may provide a vehicle that is lighter in weight than certain other electric drives, it will also have to compete with those that are not all-electric. The authors present no figures or comparisons with other methods of driving so far as weight is concerned. In spite of the intelligent use of material apparent in this design, we believe that its weight per horsepower will be substantially above other drives. For this reason we do not believe it will reduce unit costs by reducing the empty machine weight.

2. Increase the size of the machine to carry more payload. It is fairly well recognized in scraper operations that the larger the machine, the lower the unit costs, all other things being equal. Since this drive is intended for around 110 tons gross load and is, therefore, designed for a much larger machine than is produced by most manufacturers, the size factor will tend to favor it in any simple comparison of unit costs. The use of electric drives to transmit all the horsepower may definitely show to advantage where larger vehicles would be practical, but dependable components other than electric have not been developed.

The controls and auxiliary equipment that must accompany the items shown haven't been described, but we must include them also in any considerations of weight or expense.

We would expect initial costs for this system to be higher than for the other drives generally being used. In those cases where the same size shovels are available in either electric or mechanical versions, we find that the price for electric is in the range of 20–30% higher than mechanical.

We would not necessarily expect the pattern of relative costs between electric and other systems to remain the same for typical off-highway vehicles as for shovels. Two factors that are apparently different are the ratio of horsepower to machine weight and the range of speeds over which almost full horsepower is required. On the other hand, the electric drive presented is supposed to be simpler, and therefore, probably less costly than previous arrangements. There may be other compensating factors which make the weight and cost of the electric transmission system acceptable to the customer.

We believe that electric drives will certainly have a place in off-highway equipment, especially in the largest machines. It remains to be seen how well the electric will be able to compete in the larger sizes that may become popular, where other dependable transmission systems should also be available.

Authors' Closure To Discussion

NVENTIONS DISCLOSING electric drive for vehicles date back into the 1800's. The use of electric motors as propulsion equipment is by no means new, as evidenced by the early advent of the electric locomotive. The authors' intent was merely to point out that General Electric's first "motor in a wheel" was produced on Byrd's "Snowcruiser," and that such background in this and locomotive traction fields led to the present motorized wheel. This current design contains patented features which make it unique in application of electric traction to off-highway vehicles.

The authors agree with Mr. Hyler that trolley applications are a natural for electrically powered off-highway vehicles. The low cost power thus available is a boon to truck designers, enabling high-horsepower, high-performance trucks to be built with a very favorable gross to tare weight ratio.

Mr. Hyler expresses concern over the large diameter seals used in the design. The designers recognized the importance of this item, and an extensive test program was pursued to check seal performance both in the laboratory and in the field. To date, this seal arrangement has given a good account of itself.

As for parking, a simple hydraulic accumulator would provide brake assurance for the majority of cases. If a vehicle is to be left parked for extended periods, the prudent owner would block the wheels of his truck, regardless of transmission or brake type.

In regard to maneuverability, the large size of a vehicle utilizing these motorized wheels would necessitate the tractor-trailer-type design. By this, is meant the 2-wheel tractor, 2-wheel trailer, wagon-steered type of vehicle which Mr. Hyler agrees is the most maneuverable wheeled vehicle to date. Automotive steering on a tractor cannot approach the maneuverability of a 2-wheel wagon-type steered tractor.

Mr. Hyler mentions that the weight per horsepower of this drive will be substantially above that of other drives. It is difficult to see how this conclusion is reached as no vehicle design is presented in the paper. Yet it is the vehicle design which will show to advantage in regard to weight, when using the motorized wheel. The electric transmission allows the vehicle designer considerably more latitude in structure layout. As there are no critical line-up problems with drive lines, universal joints, differentials, and the like, the vehicle frame can be made more flexible (and lighter) than with conventional transmission designs.

In reply to the comments concerning the dynamic braking capacity of the motors shown in Fig. 18, it should be pointed out again that the peak continuous braking horsepower at 14 mph is in excess of 1700 hp for the four motors on the vehicle considered. At high speeds, braking horsepower is reduced because the capacity of the motor is limited by its ability to commutate high current. Motor current is held to a safe level at these speeds by a relatively simple electrical control circuit that makes it necessary to accept some reduction in braking capacity in the middle speed range. In view of the considerable braking capacity still available from the electric motors, it was felt that this braking control would be an acceptable compromise. In addition, the mechanical disc brakes provide 60,000 lb of braking effort which is available to the operator should it be required to maintain control of the vehicle.

With regard to the comments on the economics of electric drive — cost figures on mechanically or hydraulically driven equipment, such as shovels, oil well drilling rigs, passenger buses, locomotives, and the like, invariably show higher maintenance expense levels than similar equipments employing properly designed electric drive. The demonstrated reduction in maintenance expense and corresponding increase in reliability and total equipment life are the basic reasons why users willingly pay a premium for electric drive shovels.

In the case of motorized wheel drive vehicles, the first cost premium itself is very small. Economic studies of specific applications comparing similar sized diesel-electric and diesel-hydraulic drive trucks indicate savings for the electric drive units of 10–30%.

Hydrostatic Transmissions

P. C. Mortenson
Vickers Incorporated

THE NEED FOR transmitting power has become a dominant factor in the design and development of the industrial and agricultural machinery that forms a significant part of today's mechanized economy. Ideally, a device that transmits mechanical power should be capable of efficiently converting both torque and speed in an infinite number of combinations, and to do so with a smooth, uninterrupted flow of power.

Many forms of electrical, mechanical, and hydraulic transmissions have been conceived that would approach the acme of perfection. Hydrostatic transmissions, one of the two basic forms of fluid transmissions, are of immediate interest in this paper; and, aside from a brief comparison between the two basic types, our remarks will be limited to the hydrostatic type.

The most popular and well-known form of fluid transmission incorporates either a fluid coupling or a torque converter. Both of these are of the "hydrokinetic type" and transmit power by a change in the fluid velocity. This change in velocity results in a transfer of kinetic energy, thus the term "hydrokinetic." Many commercial vehicles, and most private automotive vehicles, incorporate transmissions of this type, so I am sure that you are well acquainted with the general design and performance characteristics.

The hydrostatic type transmits power by means of fluid under pressure acting upon a moving piston. Energy is imparted to the fluid by a positive displacement pump and is transmitted to a positive displacement motor, where it is converted into useful work. So, the term "hydrostatic" is used.

Fig. 1 is a schematic diagram of a hydrostatic transmission. The positive displacement pump is shown on the left. This is the unit that is driven by the engine, or by any other type of prime mover. This portion of the transmission is identified as the "A" end, or the power input end.

The fluid leaving the "A" end is delivered under high pressure to the hydraulic motor shown on the right. This motor, called the "B" end of this system, delivers power through its output shaft.

The cylinder blocks at either the "A" or "B" ends can be swung through an angle with respect to their shafts, as indicated by the dual directional arrows. This movement produces variable speed and torque from the "B" end. This will be discussed in detail later in this paper.

Since the turn of the century, design engineers have been fascinated by the characteristics of hydrostatic transmissions. Indeed, there are many reasons for this fascination, some of which are summarized as follows:

1. The input end of the transmission can be designed to deliver an infinitely variable quantity of

THE HYDROSTATIC TRANSMISSION offers optimum performance characteristics — including smooth control of shaft speed, reversibility, high efficiency, precise control of vehicle speed and travel, rapid acceleration and deceleration, and flexibility of installation.

However, these advantages are offset by several disadvantages — noise, the requirement for precision-type materials, weight and bulk, and, more important, high cost.

The author believes that the hydrostatic transmission is the answer to exact transmission performance requirements. As such demand increases, manufacturers will be able to solve the more serious problems of this transmission.*

* Paper presented at SAE National Farm, Construction, and Industrial Machinery Meeting, Milwaukee, Sept. 16, 1959.

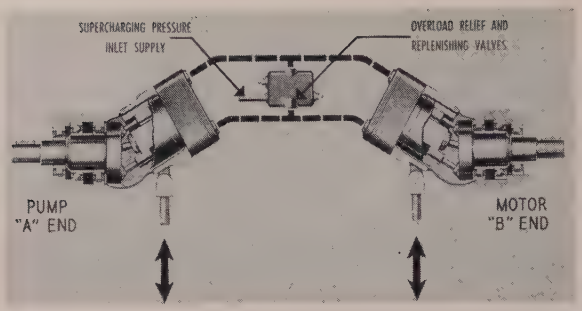


Fig. 1 — Schematic diagram of hydrostatic transmission

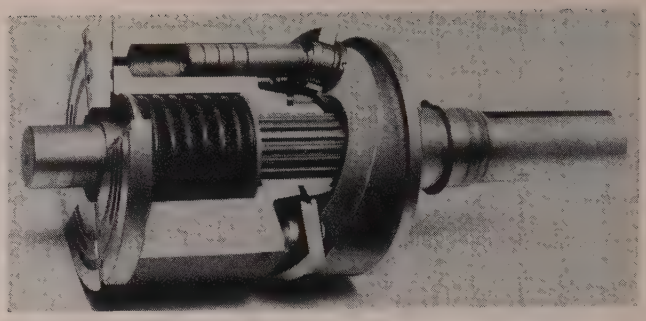


Fig. 3 — Axial piston, constant-displacement pump

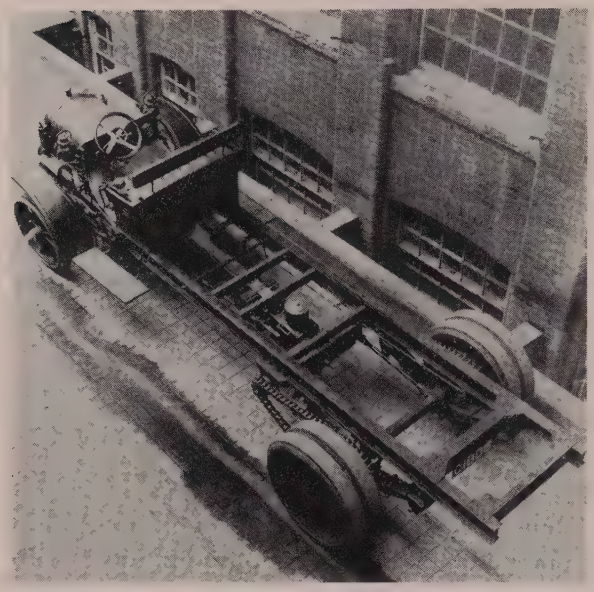


Fig. 2 — White Motor Co. truck with hydrostatic transmission, 1908

fluid. This variable quantity of fluid produces variable output shaft speed that can be readily controlled from the design maximum speed to zero speed in truly a stepless manner. The change in speed is absolutely smooth.

2. A hydrostatic transmission is fully reversible from the full design rpm in one direction, down to zero rpm, and then to full design rpm in the other direction. This reversibility is also completely smooth.

3. It is an accepted fact that other types of fluid transmissions, more particularly the fluid coupling and the torque converter types, can operate at high overall efficiency only when the ratio of the output shaft speed to the input shaft speed approaches 1/1. However, in vehicles that are called upon to maneuver frequently, the hydrostatic type of transmission has higher efficiency over a wider range of operating conditions.

4. A hydrostatic transmission does not creep, and it has no tendency to creep.

5. Extremely precise control in both vehicle speed and vehicle travel is easily obtainable.

6. The pressure relief valve is a basic part of the hydrostatic transmission circuit, and provides complete overload protection to the engine and to any auxiliary gears or shafts that are in the power train.

7. The hydrostatic transmission can accelerate or decelerate faster than any other type of power-transmitting system. Vickers has marketed a 50-hp hydrostatic transmission for aircraft applications that is capable of reversing from 8600 rpm clockwise rotation to 8600 rpm counterclockwise rotation, 5 times a sec.

8. Hydrostatic transmissions provide the designer with an unusual degree of flexibility of installation. A designer can put together an infinite number of transmission combinations by selecting units from a large range of pump and motor sizes of either fixed or variable displacement, combined with numerous types of controls.

9. Dynamic braking of the vehicle is an inherent feature of a hydrostatic transmission. Tests conducted on numerous off-highway vehicles incorporating hydrostatic transmissions proved conclusively that the conventional shoe- or disc-type brakes could be eliminated, except as required to meet specific safety codes.

10. If a variable displacement pump is used at the "A" end of the transmission and a fixed displacement motor used at the "B" end, a transmission is available that can deliver constant torque throughout the full speed range. Such a transmission can easily be matched to any normal internal-combustion engine characteristic.

11. Now, if we interchange these units and use a fixed displacement pump at the "A" end of the transmission and a variable displacement motor at the "B" end, the transmission is capable of delivering constant horsepower independent of the output speed. The use of an arrangement of this type would permit the vehicle designer to operate his engine at a governed speed, putting constant horsepower into the transmission.

12. Throughout the entire torque range, a vehicle incorporating this type of transmission can operate at a constant seed, independent of whether it is moving uphill, downhill, pulling light loads, or pulling heavy loads. If a drawbar load is suddenly removed, the vehicle speed remains practically unchanged.

The characteristics outlined above point out, to a limited degree, the reasons why both vehicle designers and operators have been interested in hydrostatic transmissions for these many years. But there are also some unresolved problems in hydrostatic transmissions that we cannot overlook. These may be summarized as follows:

1. At the present time hydrostatic transmission noise can become a problem when the units are re-

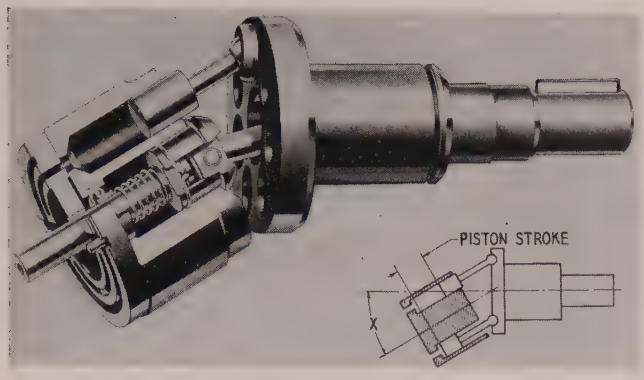


Fig. 4 — Angular-type pump

quired to operate over wide speed and pressure ranges.

2. To obtain high efficiencies, precision-type materials and workmanship are required.

3. Hydrostatic transmissions may be heavier in weight and larger in bulk than mechanical transmissions for a given horsepower rating.

4. As of today, there are no high-performance hydrostatic transmissions commercially available that can compete directly with other forms of mechanical transmissions on a dollar-per-horsepower basis. The cost problem is resolved, however, when the advantages offered justify its higher cost. Designers continue to place greater demands on transmission performance, and we are approaching the time when the hydrostatic transmission will be more nearly competitive with other transmissions. Naturally, substantial cost reductions will be made when high-production tooling and manufacturing methods can be used.

Hydrostatic transmissions were used in relatively crude form prior to the turn of this century. These units consisted of a group of conventional cylinders arranged radially around an output shaft. Severe speed limitations were inherent in this design. And this limitation, along with the large diameter requirements, limited its acceptance to large, low-speed applications.

At the turn of the century the Waterbury Tool Co., now a division of Vickers Incorporated, developed hydrostatic transmissions incorporating pistons that were disposed axially around the input and output shafts. This early design was used by the United States Navy and also by commercial ships for operating steering gears, windlasses, cargo winches, and other similar equipment. Shortly after the success of these applications was assured, the application of hydrostatic transmissions was undertaken in industry for wire rope and cable-stranding machines, for paper mill drives, steel mill drives, and for conveyor drives. Fig. 2 shows a White Motor Co. truck that was built in 1908, which had an early type hydrostatic transmission. This truck was used to haul machinery over mountain terrain with considerable success for about 6 years. It met a disastrous end in 1914 when it ran off the road and was demolished. Following those early pioneering days, the axial position type of hydrostatic transmission has been steadily improved in design concept, and today is quite popular in the heavy machinery fields.

There are several types of fluid pumps and motors that can form hydrostatic transmissions, all of which are of the positive displacement type. These can use gears, vanes, or pistons as displacement elements; but, of these, the transmissions that are required to operate at relatively high levels of efficiency must be of the piston type. At the present stage of development, this is the only tpye that can operate at pressures of 5000 psi, that has high volumetric and mechanical efficiency when new, and that retains these high efficiencies over extended periods of time. Another advantage inherent in the piston-type pump or motor is that the displacement can easily be made infinitely variable.

Fig. 3 shows a typical axial piston, constant displacement pump that can be used for either the "A" end or the "B" end of a hydrostatic transmission. Note that the pistons and connecting rods are placed around the input shaft centerline and are parallel to that centerline. Each piston is fitted into a cylindrical chamber that is an integral part of a cylinder block. A swash plate is placed concentric with the shaft centerline, but tilted at an angle to that centerline; and, as the pistons and cylinder block rotate, the pistons are moved axially in the cylinder block bores. By means of suitable valving, oil enters the rotating cylinder block when the chamber above the piston head is increasing in volume. As the piston and cylinder block further rotate, oil is forced through suitable valving out of the cylinder block head.

Fig. 4 shows a pump that is of the angular type. This pump consists basically of a drive shaft to rotate the pistons, a cylinder block to house the pistons, and a stationary valving surface facing the cylinder block bores which ports the inlet and outlet flow. The drive shaft axis is at an angle to the cylinder block axis, and rotation of the drive shaft causes rotation of the pistons and the cylinder block. Since the pistons' plane of rotation is at an angle to the valving surface plane, the distance between any one of the pistons and the valving surface is continually changing during rotation. Each individual piston moves away from the valving surface during one half of the shaft revolution, and toward the valving surface during the other half. The valve surface is so ported that its inlet passage is open to the cylinder bores in that part of the revolution where the pistons are moving away, and its outlet passage is open to the cylinder bores in that part of the revolution where the pistons are moving toward the valving surface. Therefore, during pump rotation, the pistons draw liquid into their respective cylinder bores through the inlet chamber and force it out through the outlet chambers.

A piston motor, which for all practical purposes is identical to the pump described above, receives the fluid under pressure from the pump. The motor converts fluid energy into mechanical energy in terms of speed and torque applied to the output shaft.

All of these types of piston units can be readily converted from the fixed displacement type to the variable displacement type. The axial piston swash plate can be mounted on a rotatable trunnion to vary the angle between it and the drive shaft. As the swash plate angle changes, the stroke of the pistons changes, resulting in the delivery of a variable quantity of fluid.

Fig. 5 shows the design arrangement of a variable

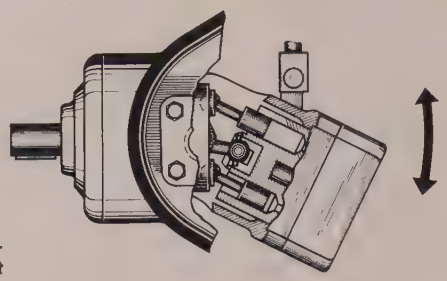


Fig. 5 — Variable-angle-type piston unit

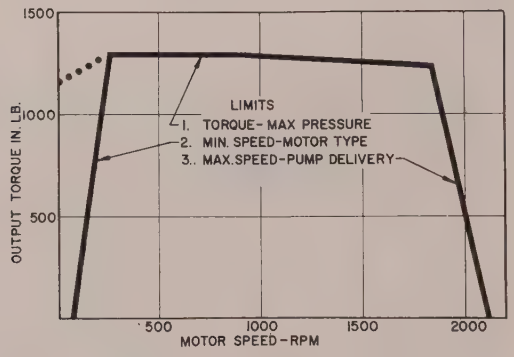


Fig. 6 — Potential output of hydrostatic transmission with variable-displacement pump and fixed-displacement motor

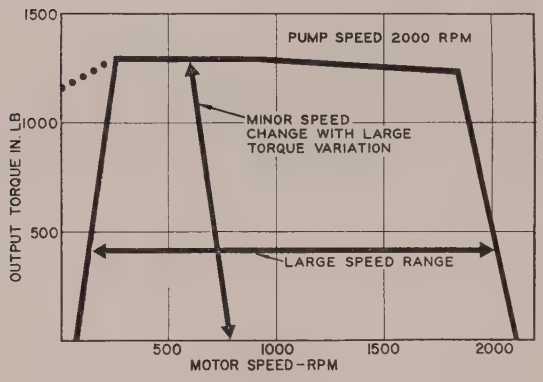


Fig. 7 — Output speed ranges of hydrostatic transmission

angle type piston unit. A yoke is arranged to permit the cylinder block to be displaced through a prescribed angle in relation to the inlet shaft. As the cylinder block is displaced angularly, the pistons are also displaced, and the stroke varies with the magnitude of this displacement. The variable delivery output is proportional to this angle through which the cylinder block and pistons are displaced.

In addition to the piston-type units described above, there are numerous types of gear and vane pumps and motors on the market that are used at the "A" and "B" ends of hydrostatic transmissions. Units of this design are more economical initially, but operate at overall efficiencies that are appreciably lower than piston-type units. Their use is limited to low-performance type of systems where efficiency is not a paramount requirement.

Designers of transmissions must give detail consideration to the general and to the specific performance characteristics of the equipment on which the transmission will be used. With the many types

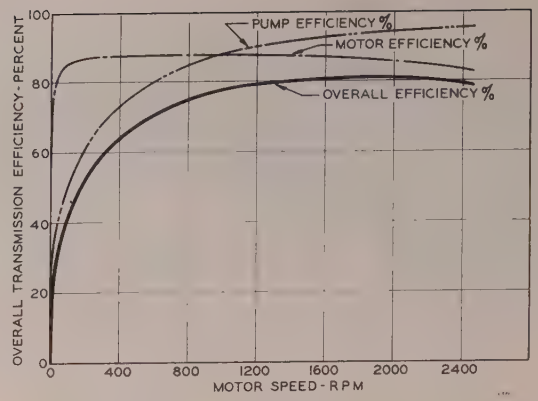


Fig. 8 — Efficiency characteristics of hydrostatic transmission for farm tractor (constant 75 hp input)

of power transmissions available today, this becomes a rather complex subject and requires a great amount of performance investigation. As far as hydrostatic transmissions are concerned, there are an infinite number of available performance characteristics, so it is impossible to describe and discuss all of the possibilities. Only a limited number of operational characteristics will be discussed here.

Fig. 6 shows the potential output of a typical hydrostatic transmission consisting of a variable displacement pump driving a fixed displacement motor. This figure uses an ordinate of transmission output torque vertically, and of motor speed horizontally. This figure is for a specific hydrostatic transmission drive that operates on a lift truck at 5000 psi. The vehicle engine speed is held constant at 2000 rpm. The coordinates are equivalent to the tractive effort and vehicle speed performance characteristics that are commonly used on mobile equipment.

The boundaries of this diagram are established as follows:

1. Along the top, the output torque is established by the proper selection of the motor size and of the operating pressure.

2. Most hydrostatic motors cannot operate consistently at speeds that approach zero rpm, and the location of the slope at the left-hand boundary is established by the low-speed characteristics of a specific motor design.

3. The maximum speed, shown as the right-hand boundary, is determined by the quantity of fluid delivered to the motor.

The significance of this design is that the hydrostatic transmission is capable of operating at any point within the boundaries just described. Any combination of output torque and motor speed within these boundaries can be transmitted in a smooth and stepless manner.

Fig. 7 points out another principal characteristic of the hydrostatic drive. The heavy horizontal line describes the large motor output speed range that is available even with a constant pump input speed. If ample engine power is available, there is no variation in torque output. Another attractive characteristic is inherent with hydrostatic transmissions. The nearly vertical line shows the minor change in motor speed as the torque changes from zero to the maximum output torque. A vehicle with this type of transmission can operate over ruts or obstacles, or up and down grades, without a noticeable change

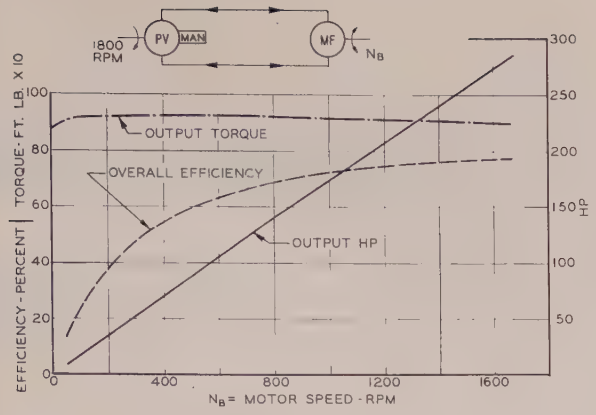


Fig. 9 — Typical performance characteristics of variable displacement pump — fixed displacement motor

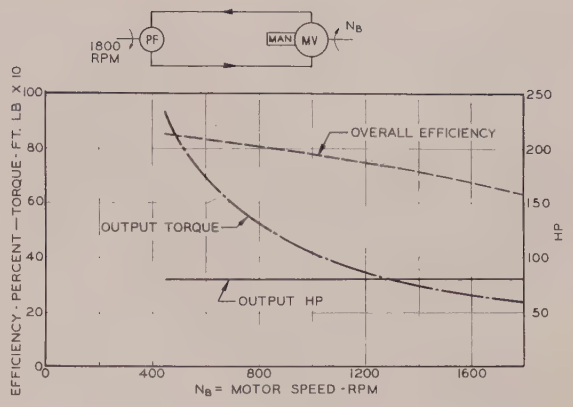


Fig. 10 — Typical performance characteristics of fixed displacement pump — variable displacement motor

in speed. The vehicle performance characteristics with this type drive are indeed excellent.

Efficiency is a transmission performance characteristic that is of considerable concern to the vehicle designer. We will discuss the efficiency characteristics of hydrostatic transmissions in terms that can be readily comprehended.

When speaking of efficiency, we must realize that "overall efficiency" is of specific interest. Throughout this paper the word "efficiency" means overall efficiency and includes all losses due to fluid flow, surface friction, ball bearing windage, fluid internal leakage, and fluid compression losses. Efficiency here is obtained by dividing the net mechanical power output of the transmission by the gross mechanical power input to the transmission.

Fig. 8 consists of three efficiency curves that make up a part of the overall operating characteristics of a specific transmission. The curves shown are for a hydrostatic transmission that was designed for use in an agricultural tractor. It was rated at 75-hp input and consisted of one variable displacement pump at the power input end and two fixed displacement motors at the power output end. The two motors provided differential action of the driving wheels. The solid line of Fig. 8 is typical of the overall efficiency that could be achieved for this installation.

No general statement of efficiency can be made without detailed description of the system and its desired operating characteristics. However, to answer this question partially, efficiency curves for two more typical transmissions will also be shown.

One of the simplest forms of hydrostatic drives makes use of a variable displacement pump and a fixed displacement motor. The performance characteristics of this type of drive are shown on Fig. 9. Note that this arrangement produces a torque output that is constant throughout the entire speed range, and a horsepower output that is directly proportional to speed. The efficiency curve shown is for a pump operating at its maximum rated speed and pressure. These characteristics will alter somewhat with changes in speed and pressure. The practically constant torque characteristic from starting speed to maximum speed is of interest to designers of many types of agricultural and industrial machinery.

When a fixed pump drives a variable displacement motor, a transmission is available that is capable of providing constant horsepower output—even though the speed and torque on the output shaft may vary over relatively wide ranges. Fig. 10 presents the typical performance chcaracteristics of a transmission of this type. To be consistent, the motor size selected for the data shown in this figure is identical with the motor size used in Fig. 9. This type of constant horsepower transmission is of specific interest to designers of equipment that use governed engines, and who want to operate their machinery continuously at maximum horsepower output to obtain the maximum work capability. This transmission has the ideal theoretical characteristics of dividing torque and speed into an infinite number of combinations to use the power from a given engine most effectively. However, practical limitations are imposed by the torque range of the motor, by the relatively larger motor size, and by generally lower efficiency.

Conclusions

In the preceding discussions we have considered two basic types of hydrostatic transmissions. The first consisted of a variable pump and fixed motor, and the second of a fixed pump and a variable motor. It is also possible to use two variable displacement units to form a hydrostatic transmission. Such a combination is capable of delivering either constant torque or variable torque from the output shaft. By proper control of the "A" and "B" ends of this type of transmission, an infinite number of torque-speed relationships are available. Transmissions of this type can theoretically provide the designer with an unlimited number of power output performance characteristics.

We can say that the hydrostatic transmission has provided, and will continue to provide, the optimum transmission performance characteristics, but that to date the relatively high cost continues as the major unsolved problem. As our design and manufacturing skills increase, and as transmission performance demands become more exacting, solutions to this problem become more probable. As soon as this has been accomplished, the hydrostatic transmission will become preeminent in the agricultural and industrial machinery field.

Discussion of this paper will be found on p. 256, following the Bowers paper.

Hydrostatic for

THE HYDROSTATIC DRIVE has ceased to be a "promising development," the author of this paper believes. The transmission has become a proved system, offering unique advantages to the designers and users of construction machines and others who place a premium on maneuverability, controllability, and the full availability of maximum power with changing vehicle speed.

The author discusses the advances made to date, various applications of the transmission (including tractor and shovels), and performance curves.*

E. H. Bowers
Dowty Hydraulic Units Limited

VEHICLE TRANSMISSIONS are required primarily to provide a torque and speed variation between the prime mover and driven wheels or tracks, to change the sense of rotation of the drive line for reverse movement, to allow very low speeds of vehicle movement which are beyond the range of prime mover speed variation, and to permit the vehicle to start smoothly from rest.

Ideally, the transmission should provide a stepless range of ratios for forward and reverse vehicle movement, which can be changed at any desirable rate under any load conditions, and to operate with negligible power losses. Fig. 1 shows the type of performance obtained with a mechanical change speed box compared to the performance of an infinitely variable drive over the same range of vehicle speeds. It may be noted that the maximum torque developed by the infinitely variable transmission is limited by the onset of wheel or track slip, but that above the speed at which this occurs, the full horsepower of the prime mover can be fed through for traction purposes without interruption. The elimination of the engine speed change steps, inherent with the mechanical change speed box, means that the ideal ratio for any vehicle speed can be obtained and the work output of a vehicle with a given power can be increased if the efficiency of the infinitely variable drive is high.

Fig. 1 also illustrates one useful method of defining the performance of a stepless transmission — by using the "full power torque/speed ratio." Although a drive of this type is usually capable of operating from an infinite speed ratio for starting purposes up to a limiting speed ratio, the performance curve is divided into two parts: the first part from zero output speed gives constant torque at rising horsepower, the second part gives falling torque at constant power. The ratio of the constant torque speed range to the total speed range is 0.25/1 in Fig. 1, or quoting the output first, is usually stated as "4/1 torque/speed ratio." The significance of this will be discussed later.

There are many other important requirements concerning the control of ratio change, braking, and limitations on bulk and weight, but they vary considerably between different types of vehicles and cannot be included in a basic definition of requirements.

Types of Hydrostatic Transmission

Before considering how the various possible arrangements of hydrostatic transmissions measure up to these requirements, some definition of a hydrostatic transmission is necessary.

Basically, a hydrostatic transmission is a hydraulic circuit embodying two positive displacement hydraulic units interconnected so that mechanical power input to one unit may be transmitted by a flow of fluid under pressure to a second unit which translates the hydraulic power back into mechanical power. For uninterrupted power transmission this means in practice a rotary driven hydraulic pump supplying fluid under pressure to a rotary hydraulic motor, though the roles of pump and motor may be reversed if an input torque is applied to the unit normally producing output torque.

The simplest possible form of hydrostatic transmission, using two interconnected fixed positive displacement units, offers little advantage for vehicle transmission purposes. No variation in torque ratio between input and output is possible and variation of speed can be achieved only below the fixed ratio by spilling fluid under pressure, a power-wasting procedure analogous to a slipping clutch. If, however, one or both of the units have a variable displacement, an infinitely variable ratio control can be obtained.

Very many different arrangements are possible to meet a wide range of performance requirements for a variety of vehicles. Each of these arrangements accentuates some aspects of the performance at the expense of other aspects, and a brief analysis of

* Paper presented at SAE National Farm, Construction, and Industrial Machinery Meeting, Milwaukee, Sept. 16, 1959.

Transmissions

Vehicles

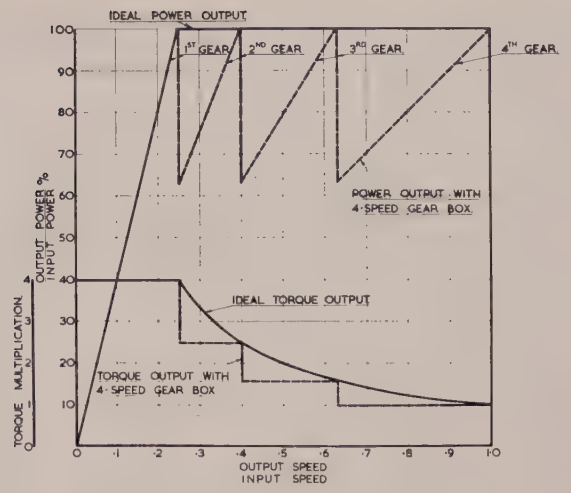


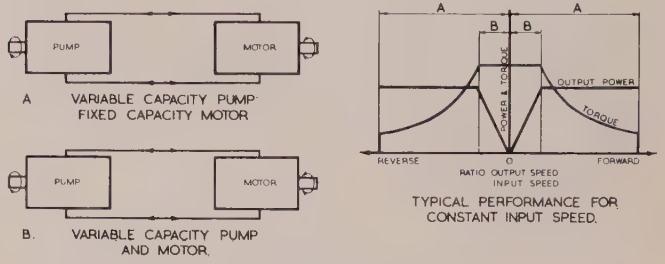
Fig. 1 — Torque and power output of ideal stepped and step-less transmissions

these basic arrangements is worthwhile as it shows the wide range of application.

There are three well-defined groups of arrangements: (1) the straight hydrostatic transmission with no mechanical connection between the input and output units; (2) a group which may be called "case reaction" where input torque is directly fed through to the output; and (3) a group of arrangements employing single and double mechanical differential gears as a supplementary drive path.[1, 2, 3]

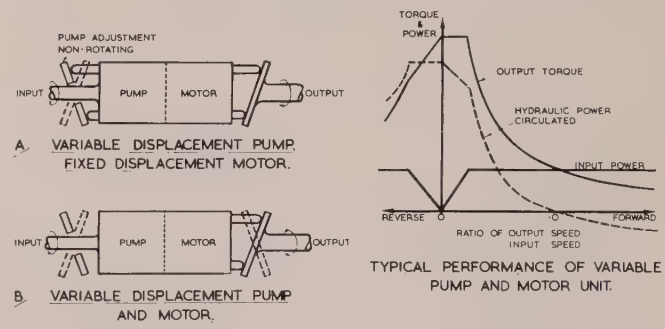
Figs. 2–5 show diagrammatically the most straightforward arrangement in each of these groups. Group 1 are suitable for use with more than one output motor, can eliminate the need for a mechanical transmission entirely, and give an equal performance in forward and reverse speeds. The input pump in Group 1 can feed power to one or more hydraulic motors which can be situated on or inside the wheels or tracks and may develop the required output torque without reduction gears. Groups 2 and 3 can have only one output shaft, can be locked up in a fully mechanical drive in some cases, but have a limited reverse rotation performance. Groups 2 and 3 can be regarded, in fact, as infinitely variable change speed boxes better suited to those vehicles which do not require frequent reversals of direction, do not need multiaxle drive, and cannot benefit from the elimination of the mechanical drive line, a field in which the torque converter is well established.

Group 1 transmissions are well-suited to the majority of farm and construction vehicles, and the remainder of this paper deals almost exclusively with this group. This does not mean that the other groups are not suited to these vehicles, but in a short paper a full discussion of all types cannot be presented and it seems logical to suppose that if a hydrostatic transmission that transmits all the power can be produced with a satisfactory size, weight, and efficiency, the types that transmit only



1. EQUAL PERFORMANCE WITH EITHER DIRECTION OF OUTPUT ROTATION
2. STARTS FROM ZERO OUTPUT SPEED WITHOUT POWER WASTAGE
3. TORQUE SPEED RATIO A/B: OPTIMUM WITH FIXED CAPACITY MOTOR PROBABLY BETWEEN 4:1 & 5:1 WITH PRESENTLY AVAILABLE UNITS
4. OUTPUT SPEED THEORETICALLY INFINITE FOR VARIABLE PUMP-VARIABLE MOTOR COMBINATION; LIMITED PRACTICALLY BY MAXIMUM EFFICIENT MOTOR SPEED
5. HYDRAULIC UNITS MAY BE REMOTELY MOUNTED FROM EACH OTHER ONLY CONNECTING MEDIUM IS PIPING.
6. CONTROLS ON PUMP UNIT ONLY FOR SYSTEM A; REQUIRED ON BOTH UNITS FOR SYSTEM B
7. BOTH SYSTEMS SUITABLE FOR MULTI-MOTOR CIRCUITS.

Fig. 2 — Group I — straight hydrostatic drives



1. UNLESS PUMP HAS GREATER DISPLACEMENT THAN MOTOR, VARIABLE MOTOR NEEDED FOR REVERSE
2. EXCESSIVE CIRCULATION OF HYDRAULIC POWER LIMITS REVERSE PERFORMANCE.
3. MAXIMUM EFFICIENCY AT 1:1 SPEED RATIO.
4. OVERDRIVE OF 2:1 WITH EQUAL DISPLACEMENT OF PUMP & MOTOR UNITS
5. SYSTEMS A & B START FROM ZERO SPEED WITHOUT POWER WASTAGE
6. SYSTEM WITH FIXED PUMP CAPACITY & VARIABLE MOTOR WILL NOT START FROM ZERO WITHOUT POWER WASTAGE. IF MAXIMUM MOTOR CAPACITY IS SAME AS PUMP CAPACITY, MINIMUM SPEED RATIO IS 0·5 WITHOUT SLIPPING FLUID FROM PUMP
7. UNITS MUST BE MOUNTED TOGETHER.

Fig. 3 — Group II — mechanical interconnection (case-reaction)

[1] "Differential Type Hydrostatic Transmissions," by C. L. Sadler, Seventh National Conference on Industrial Hydraulics, Vol. 5, 1951, pp. 41–50.
[2] "Power Dividing Transmissions," by D. L. Bedingfield. Engineering, Special Publication, 1958.
[3] "M.I.R.A. Positive Displacement Hydraulic Transmission." *The Engineer*, Vol. 204, 1957, p. 568.

a proportion of the power by hydraulic means can certainly be designed to give an acceptable performance.

Requirements of Hydrostatic Units

Hydrostatic drives are not new. They have been in use industrially and in ships for many years, and, as has been pointed out in previous papers,[4,5,6] the hydrostatic drive has almost ideal performance characteristics for construction vehicles if an acceptable overall efficiency can be obtained. The reason why hydrostatic drives have not been available for vehicle use in the past is simply that hydraulic units with an acceptable efficiency have not been available within the bulk and weight limitations which the vehicle designer imposes. In recent years this situation has changed and the number of vehicles using hydrostatic transmissions for propulsion is now growing rapidly. This change has been brought about by the development of pumps and motors specifically for transmission purposes and in the case of my own company, specifically for vehicle transmissions. The three main requirements for vehicle transmission units are:

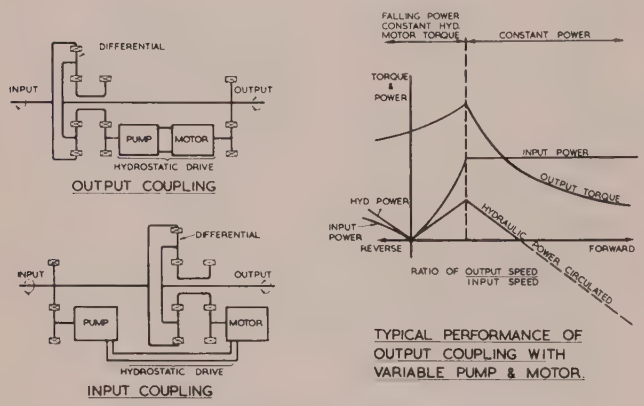
1. A high overall efficiency over a wide range of speed ratios and operating conditions
2. Input speeds high enough to permit direct coupling to the prime movers used in vehicles
3. A size, weight, and cost which are acceptable to the vehicle designer.

Efficiency comes first on the list since the worth of any new form of transmission is initially judged to be inversely proportional to the power losses. In view of the number of high-efficiency hydraulic pumps and motors already developed, it might be thought that this problem was already solved. However, because of the peculiar operating conditions in an infinitely variable transmission, a conventional high efficiency will not produce a satisfactory transmission efficiency. The losses in a conventional hydraulic pump or motor are normally expressed as a percentage of the power obtained from the product of the maximum flow and pressure.

In a hydrostatic transmission having a torque speed ratio of 4/1 (Fig. 6), the power that can be transmitted over the constant horsepower range of speeds is not the product of the maximum flow and pressure obtainable, but this figure divided by the torque speed ratio. In this case one-fourth of the maximum theoretical rating. As the losses do not decrease in direct proportion as the power transmitted is decreased, it can be seen that a variable delivery pump that has a high efficiency under normal rating conditionns, has a much lower efficiency under transmission conditions, and that special units having exceptionally high efficiencies are essential if an acceptable transmission efficiency is to be maintained. A variable delivery pump with an overall efficiency rating of 92% under normal rating conditions will, with a motor of similar rating, provide a transmission efficiency of about 60%.

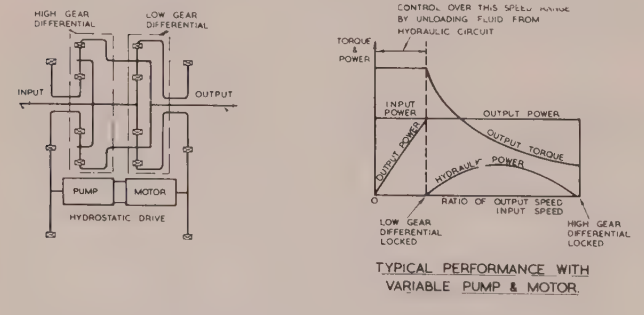
Operating speeds, size, and weight have generally been brought into line by the use of high pressures — 4000–5000 psi — and the development of designs with a high specific output. The need for high pressures and high efficiency has made the use of piston-type pumps and motors essential, though some transmissions of the differential type (where only a proportion of the power is transmitted hydraulically) have been built using vane-type displacement elements. Units for the transmission of very large powers represent some problems as the diesel or gasoline motor designer has a tendency to standardize on cylinder dimensions and to add cylinders to obtain increased power, maintaining the same operating speed. The hydraulic engineer having to use a variable stroke machine has not at the moment a similar facility. However, units capable of transmitting up to 350 hp have been in very successful use in locomotives manufactured by Antonio Badoni of Lecco in Italy for some years.[7]

Hundreds of thousands of operating hours have



1. IN ONE SPEED DIRECTION BOTH SYSTEMS WILL INCREASE TORQUE/SPEED RATIO OF STRAIGHT HYDROSTATIC TRANSMISSION UNITS FOR A GIVEN POWER; OR INCREASE POWER CAPACITY OF THE SAME HYDROSTATIC UNITS OVER A GIVEN TORQUE/SPEED RATIO.
2. REVERSE PERFORMANCE EITHER IMPOSSIBLE OR LIMITED BECAUSE OF EXCESSIVE CIRCULATION OF POWER THROUGH HYDRAULIC UNITS.
3. OUTPUT COUPLING NEEDS VARIABLE PUMP & MOTOR TO WORK WITHOUT POWER WASTAGE OVER FALLING POWER-CONSTANT HYDRAULIC MOTOR TORQUE RANGE.
4. INPUT COUPLING AVOIDS NEED OF VARIABLE MOTOR FOR CONDITION 3.

Fig. 4 — Group IIIA — single differential systems



1. VARIABLE CAPACITY PUMPS & MOTORS NEEDED TO OBTAIN LOCK-UP CONDITIONS IN LOW & HIGH GEAR DIFFERENTIAL
2. POWER WASTAGE BELOW LOCK-UP SPEED OF LOW-GEAR DIFFERENTIAL.
3. NO REVERSE, BUT OVERDRIVE POSSIBLE
4. MAXIMUM PROPORTION OF POWER TRANSMITTED HYDRAULICALLY INCREASES AS SPEED RANGE IS INCREASED E.G. FOR RANGE OF SPEED RATIOS OF 4:1 MAXIMUM HYDRAULIC POWER IS APPROXIMATELY 33⅓% OF TOTAL OUTPUT POWER

Fig. 5 — Group IIIB — double differential systems

[4] "Hydrostatic Transmission in Earthmoving Equipment," by H. V. Parsley. Paper presented at SAE National Tractor Meeting, Milwaukee, September, 1954.
[5] "Earthmoving Transmissions — Where Are We Going?," by R. B. Clark. Paper presented at SAE Central Illinois Section meeting, Peoria, Ill., May, 1956.
[6] "Trends in Hydraulic Systems," by P. C. Mortenson. Paper presented at SAE National Farm, Construction, and Industrial Machinery Meeting, Milwaukee, September, 1958.
[7] "High-Pressure Hydraulic Power Transmissions," by H. Thoma. Paper presented at Institution of Mechanical Engineers Meeting, December, 1957.

been accumulated by hydrostatic transmissions in vehicles over a range of powers and duties. The rugged trouble-free performance over a long life which has been characteristic of these drives in industrial use can be retained in well-designed units operating in suitable hydraulic circuits. In the writer's experience, not only can the overhaul periods demanded by the user be met with only minimum maintenance, but the slight cushioning of the hydraulic drive, even when violently handled, lengthens the life of associated drive line components in much the same way as the torque converter.

Range and Performance

The hydrostatic transmission, like all other infinitely variable transmissions, is subject to limitations. However, these are not fundamental and the range of performance will unquestionably be increased by development during the next few years.

At present, a transmission employing a variable input unit and fixed output unit, with a torque speed ratio of between 4/1 and 5/1 can give an efficiency of better than 80% over the constant power range. This is shown in Figs. 6–8. It may be noted that the highest efficiencies are obtained if the speed of the prime mover is reduced in proportion to the power to be transmitted. This function coincides with internal-combustion engine performance, as the best fuel consumption is obtained by operating the motor at near maximum torque and varying the rpm to meet the power requirements.

Higher speed ratios can be obtained with a given displacement if the output unit is also variable. But the output unit, or units, must either have a larger maximum displacement than the input units or operate at higher speeds than the input units, and the control system must extend to the output units. We may expect the development of units capable of maintaining the present standard of efficiency over wider speed ranges in the next few years. Many types of construction vehicles have two fairly distinct vehicle speed ranges, for working, and for transport, and in these cases the use of a single mechanical change speed, or hydraulic controls which change the speed range by reverting from 4-wheel to 2-wheel drive may prove equally attractive.

Advantages and Comparative Performance

So far the general philosophy of the hydrostatic transmission has been considered against a background of requirements. The real value of the transmission is, nevertheless, determined by its actual performance in different types of machines

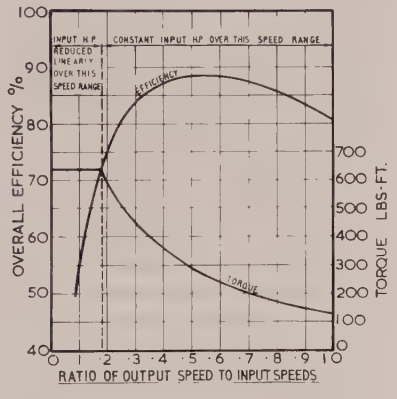

INPUT H.P. REDUCED LINEARLY OVER THIS SPEED RANGE
CONSTANT INPUT H.P OVER THIS SPEED RANGE
EFFICIENCY
TORQUE
OVERALL EFFICIENCY %
TORQUE LBS-FT.
RATIO OF OUTPUT SPEED TO INPUT SPEEDS

Fig. 6 — Performance of complete hydrostatic transmission input 45 bhp at 1500 rpm

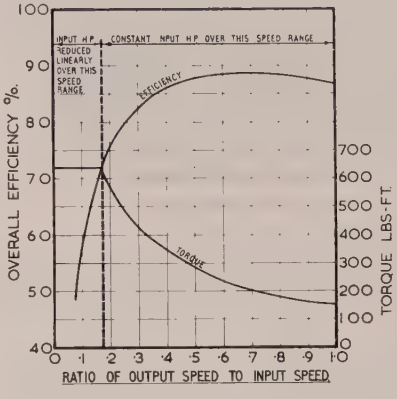

INPUT H.P. REDUCED LINEARLY OVER THIS SPEED RANGE
CONSTANT INPUT H.P OVER THIS SPEED RANGE
EFFICIENCY
TORQUE
OVERALL EFFICIENCY %.
TORQUE LBS-FT
RATIO OF OUTPUT SPEED TO INPUT SPEED

Fig. 7 — Performance of complete hydrostatic transmission input 30 bhp at 1000 rpm

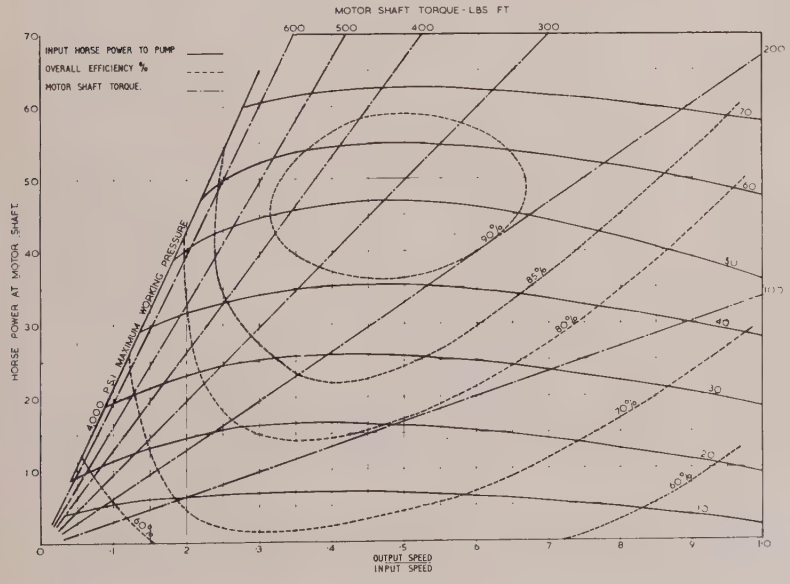

MOTOR SHAFT TORQUE - LBS FT
INPUT HORSE POWER TO PUMP
OVERALL EFFICIENCY %
MOTOR SHAFT TORQUE.
HORSE POWER AT MOTOR SHAFT
4000 P.S.I. MAXIMUM WORKING PRESSURE
OUTPUT SPEED / INPUT SPEED

Fig. 8 — Performance of complete hydrostatic transmission input speed 1750 rpm

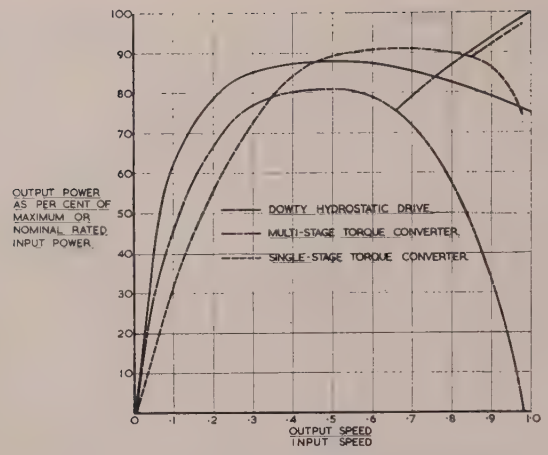

OUTPUT POWER AS PER CENT OF MAXIMUM OR NOMINAL RATED INPUT POWER
DOWTY HYDROSTATIC DRIVE
MULTI-STAGE TORQUE CONVERTER
SINGLE-STAGE TORQUE CONVERTER
OUTPUT SPEED / INPUT SPEED

Fig. 9A — Conventional comparison of hydrostatic and torque converter transmission

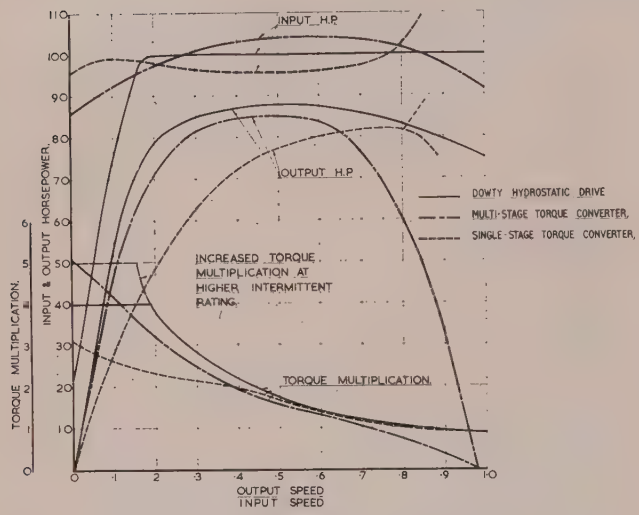


Fig. 9B — Comparison of hydrostatic and torque converter transmissions

relative to other transmissions as measured by the rate of work and cost per unit of work. However, before analyzing the application of a transmission of the type recording the performance (Figs. 6–8) to various machines, we can set down the advantages of the hydrostatic transmission and show the main differences between the performance obtained from hydrostatic and hydrokinetic or torque converter transmissions.

In the simplest terms the salient advantages of the hydrostatic transmission are as follows:

1. The ability to operate efficiently over a wide range of torque speed ratios and to reverse rotation at a controlled rate which is unaffected by the value and direction of the output load up to a limiting maximum torque.
2. The ability to remain stalled under full load without damage and with a very low power wastage.
3. The ability to hold a selected speed ratio accurately against driving or braking torques up to preset limiting values. This ability may be limited by simple automatic overriding controls where desirable.
4. The ability to produce more torque for less space and weight and with less inertia than any other type of remote actuator transmission.
5. The ability to transmit power without the vehicle layout limitations imposed by a mechanical drive line.

In comparison with the torque converter the hydrostatic transmission offers a high efficiency curve over a much wider range of speed ratios with very much lower losses in the stalled or near stalled condition, but because of the absence of a mechanical drive line it does not offer the possibility of lockup.

Fig. 9A shows the general shape of the conventional efficiency curves of two typical torque converters of 3/1 and 5/1 stalled torque ratio, plotted on the same scale as a hydrostatic transmission. Since these curves are plotted as actual output power as a percentage of a nominal or maximum power input and do not take into account variations of input power, they do not really show the relative efficiencies of the units.

If the average power inputs to the three transmissions are arbitrarily rated at 100 and the outputs rated as a percentage of the actual input horsepower under any given condition, the hydrostatic transmission which operates with the constant governed input speed shows up to greater advantage as shown in Fig. 9B. With the addition of the torque multiplication curves the full performance of the three units is given plotted against speed ratio. However, even these curves do not show the effect of the reduced input horsepower required by the hydrostatic transmission so graphically as the horsepower losses plotted in Fig. 10. In this case the same scales have been used as in Fig. 9B and the excess of input horsepower over output horsepower, which must be dissipated in heat, has been plotted against speed ratio.

Performance in Vehicles

Several different types of vehicles with hydrostatic drives have been operating for a number of years and have demonstrated clearly that this form of drive has the robustness to stand up to hard work and abuse with only trivial maintenance.

In Germany the Guldner material handling truck and a dumper have been in production for 5 years; the Badoni locomotives have been produced for 10 years in Italy. In England a prototype tractor was produced by the National Institute of Agricultural Engineering in 1954;[8] this has been followed up by the application of Dowty units to tractors, road rollers, and locomotives.

Of all these vehicles the ease of control and maneuverability by means of a single lever or pedal, which controls the direction, speed, and braking of the vehicle, is self evident. This can most easily be shown by a film and needs no further comment.

As has been previously mentioned, efficiency in the final anaylsis is a matter of cost per unit of work. The effect of the type of transmission used on the work capacity of a given machine or rate of work is important, and the ease and precision of control already mentioned plays an important part in determining the working rate of may types of machines. However, if the advantage of increased

[8] "Application of Hydrostatic Transmissions to Tractors," by H. J. Hamblin. *Journal* of Institute of British Agricultural Engineers, Vol. 12, March, 1956, p. 10.

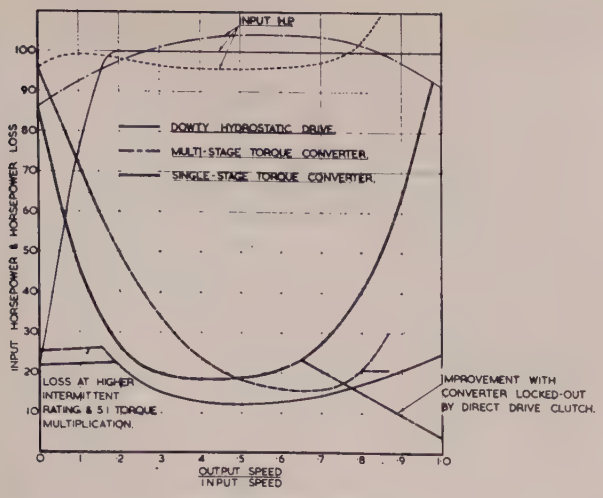

INPUT H.P
DOWTY HYDROSTATIC DRIVE
MULTI-STAGE TORQUE CONVERTER.
SINGLE-STAGE TORQUE CONVERTER.
INPUT HORSEPOWER & HORSEPOWER LOSS
LOSS AT HIGHER INTERMITTENT RATING & 5:1 TORQUE MULTIPLICATION.
IMPROVEMENT WITH CONVERTER LOCKED-OUT BY DIRECT DRIVE CLUTCH.
OUTPUT SPEED / INPUT SPEED

Fig. 10 — Comparative losses of hydrostatic and torque converter transmissions shown on Fig. 9B

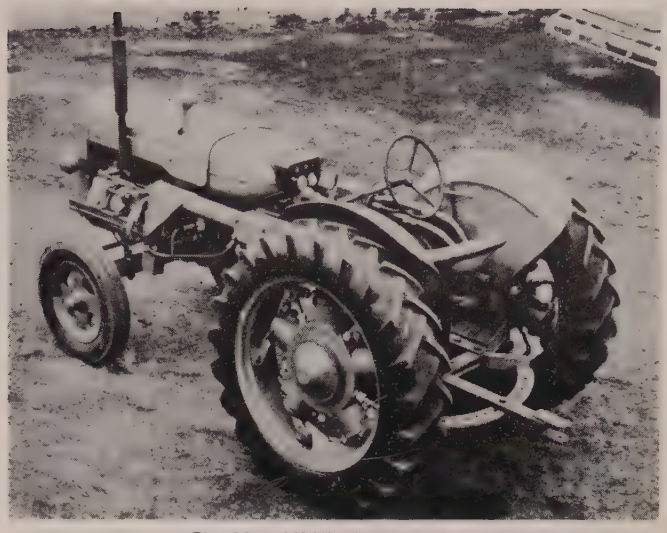


Fig. 11 — NIAE wheeled tractor

Fig. 12 — Dowty wheeled tractor

maneuverability is discounted, the efficiency of the transmission in terms of work done per unit of fuel consumption can be taken as a measure of efficiency of the drive line. This contributes to the overall efficiency of the machine, and can be assessed in machines such as tractors and locomotives which produce drawbar pull. Direct comparisons are still, however, difficult since the user of a vehicle fitted with an infinitely variable transmission almost always adjusts the speed ratio to make full use of the maximum engine horsepower under any conditions.

In the case of a tractor ploughing, the driver alters the ratio to suit as the soil conditions or gradient change, and in consequence, works faster. In the case of the locomotive, the ability to use full power continuously during acceleration permits the locomotive to handle either a heavier load in the same time cycle or the same load in a shorter time cycle, and advantage is always taken of this. In both cases, the rate of work is increased by 10–25% according to conditions. If the quantity of fuel used per acre of ploughing or per ton/mile is compared with machines having mechanical transmissions and the same engine horsepower the fuel consumption is, within fairly narrow limits, identical.

Hydrostatic Drives for Tractors and Shovels

In a short paper of this type it is not possible to discuss in detail the advantages of a hydrostatic transmission for different specialized vehicles, or to consider the various control refinements that can be introduced to simplify still further the driver's duty by integrating engine speed, vehicle speed, and load to maintain optimum operating conditions automatically. There are, however, a number of outstanding advantages that merit some mention and if three typical machines — the wheeled tractor, the crawler tractor, and the shovel loader — are considered, the effect of the hydrostatic transmission on the performance of the vehicle, and some aspects of the integration of the drive into the machine design can be briefly noted.

Fig. 11 shows the wheeled tractor of the National Institute of Agricultural Engineering (NIAE), and Fig. 12 the Dowty wheeled tractor. Both these machines have been built using the same production tractor for basic components, but they show very different approaches so far as the installation of the hydraulic equipment is concerned. The NIAE machine uses a single engine-driven pump feeding hydraulic power to two separate low-speed hydraulic motors within the wheels, which develop the maximum driving torque without reduction gearing. In contrast to this the Dowty tractor (which was built mainly to obtain performance data as nearly comparable with the standard machine as possible) is modified only to the extent of having the 6-speed gearbox removed and replaced by a hydraulic pump and motor, the latter feeding torque into an unchanged differential and reduction gear in the rear casing.

Each of these arrangements has advantages. The NIAE arrangement completely eliminates mechanical gearing and can permit an open frame design which is useful for mid-mounted agricultural implements. The Dowty arrangement minimizes the change to the tractor, retaining the front and rear portions unchanged. Further variations are possible if two high-speed hydraulic motors are used within the wheels with the necessary reduction gears. It is comparatively easy, although expensive because of the high bearing loads involved, to develop high driving torques within large wheels; but if small wheels are used the high-speed motor with reduc-

tion gearing becomes essential. Additionally, the torque produced by a high-speed motor and gearing can be changed by simple change of ratio, whereas the full torque wheel motor must be re-dimensioned.

It does not seem possible at this stage to say that one arrangement is better than the other, and it appears likely that the answer may vary for different types of machines and be determined mainly on manufacturing costs, since there is not likely to be any substantial difference in performance.

Fig. 13 shows graphically the difference between the Dowty tractor and the standard machine fitted with a 6-speed gearbox and mechanical clutch. It may be noted that the full drawbar pull from first gear cannot be realized owing to wheel slip but is probably necessary to enable starts to be made with a mechanical clutch against very heavy loads and to permit very low forward speeds, neither of which are limitations with a hydrostatic transmission.

Crawler tractors make very different demands on the hydrostatic transmission. Here the ability to provide two separate transmission lines of infinitely variable ratio to two tracks from one engine within the space limitations of the tractor hull opens up an entirely new field of crawler tractor control. With a separate speed control for each track, any desired radius of turn can be obtained and varied continuously. Braking loads on the inner track can be used regeneratively and fed through to the outer track and at no time need the drive to either track be disconnected. It is, of course, possible to make spot turns by putting one track in forward gear and the other in reverse gear.

Shovels and shovel loaders can benefit from another aspect of the hydrostatic drive performance. A conventional machine of this type is fitted with a torque converter which demands a substantial power input in the stalled condition and as the majority of the power required for bucket operation is needed when the vehicle is either stalled or nearly stalled, an engine which is capable of producing the power for bucket movement in addition to producing the power demand by the stalled torque converter must be provided. In contrast, a hydrostatic transmission can provide the same stalled torque for a very low power input, certainly less than 25% of the figure demanded by the torque converter. If an automatic control is fitted which progressively reduces the drive ratio to a figure approaching infinity as the machine stalls when the bucket is in operation, a great deal more power is released for bucket operation, or alternatively the

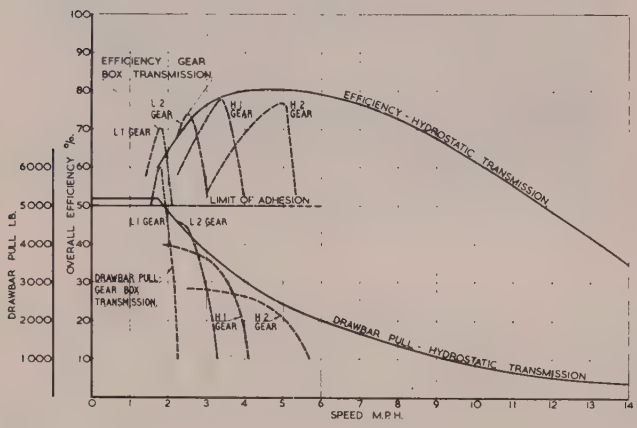


Fig. 13 — Dowty wheeled tractor performance

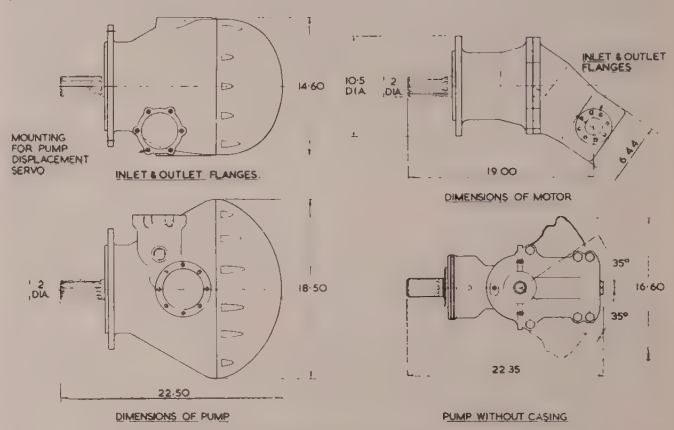


Fig. 14 — Installation dimensions of 60-hp units

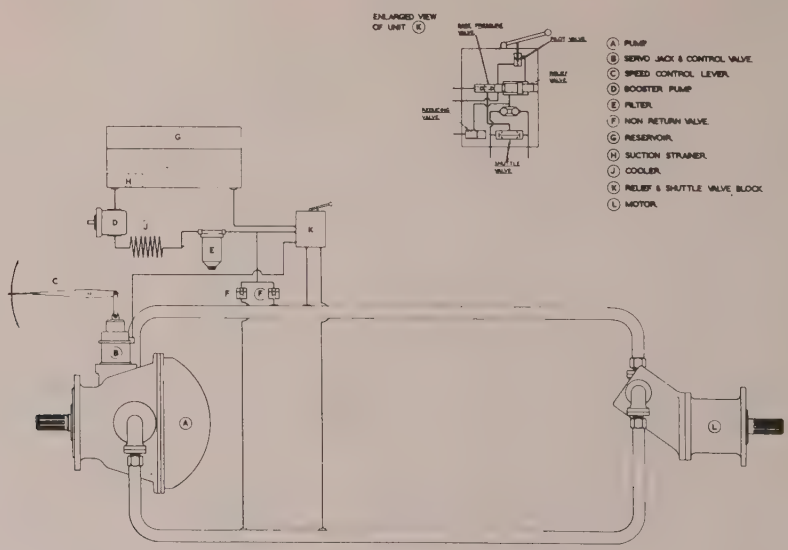


Fig. 15 — Circuit diagram

engine power may be reduced while maintaining the same performance.

Application Considerations

If we assume that a satisfactory case has been made for the performance and life of the hydrostatic transmission, its application depends upon two factors — its cost and ease of installation.

Cost is difficult to discuss without a detail design analysis, which cannot be undertaken in a general paper of this type. Present information suggests that although the hydrostatic transmission is not competitive cost-wise with a simple change speed box and mechanical clutch, it can be produced for the same or a lower price than the nearest equivalent transmission — a multispeed reversible hot shift box — if the scale of production is similar.

Reductions in cost and improved layouts can be obtained in some types of vehicles and self-powered mobile machines if the mechanical drive line is eliminated. The motive power unit may be placed across the framework parallel to the axles, or placed adjacent to one drive if several independent drives are required as in the case of road laying machinery and combine harvesters. Open frameworks between the engine and driven wheels can be arranged easily to accommodate bulky loads in such vehicles as straddle carriers, and to provide visibility in the case of agricultural and horticultural machines with midmounted implements.

Installation dimensions of typical hydrostatic units are given in Fig. 14. The particular units shown are capable of transmitting 60–70 hp with a torque speed ratio of 4/1 to 5/1. They are normally installed in a circuit with a pressure charging pump, as shown in Fig. 15.

Figs. 16 and 17 show a 30,000-lb road roller made by Marshall Sons & Co. Ltd., England, incorporating these transmission units, and the performance of the machine.

Figs. 18 and 19 show a narrow gage locomotive made by Ruston & Hornsby Limited, England, incorporating the same transmission units and the performance data. When the limit of adhesion is reached the transmission does not "run away" — the selected ratio is held with only the appropriate amount of wheel slip.

It may be noted that neither of these machines requires any special cooling provision for the hydraulic fluid under temperature operating conditions and the hydraulic fluid remains within 60 F of the ambient. This in itself demonstrates the low losses rejected in heat.

Conclusion

For nearly 50 years the hydrostatic transmission has attracted the attention of vehicle designers as a system offering the solution to many of the problems confronting them. During this period very many experimental machines have been built, but only in the last 10 years has development produced systems that can operate with an acceptable efficiency within the limitations imposed by the vehicle designer. To some extent the recent development of the hydrostatic transmission has been overshadowed by the widespread adoption of the torque converter. This is unfortunate, since while the performances of the two types of hydraulic drive do overlap, their ranges of applications are in the main different, and they are complementary one to

Fig. 16 — Marshall road roller

Fig. 18 — Ruston and Hornsby locomotive

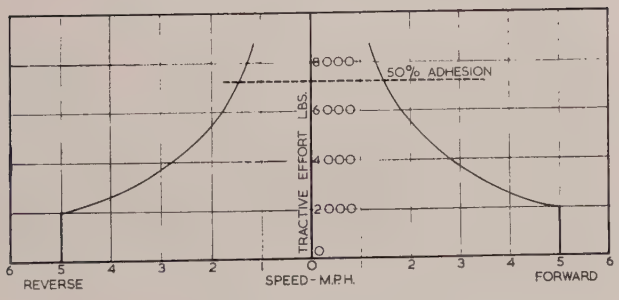


Fig. 17 — Marshall road roller performance

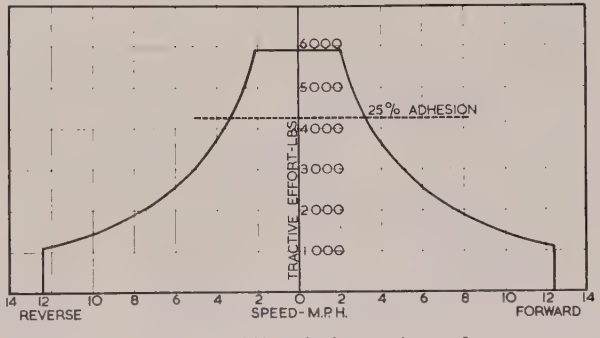


Fig. 19 — Ruston and Hornsby locomotive performance

the other rather than competitive.

In this paper the writer has attempted to show that the hydrostatic drive has ceased to be a "promising development," but has become a proved system, offering unique advantages to the designers and users of construction machines and many others, where there is a premium on maneuverability, controllability, and the full availability of maximum power with changing vehicle speed.

In the future we may expect the straight hydrostatic transmission to offer higher speed-torque ranges, further small increases in efficiency, and an increased range of higher horsepower ratings. At this stage it is difficult to predict the range of application of the differential type of hydrostatic transmissions which are almost directly competitive with the torque converter. It may well be that where very high efficiencies are required within severe weight and space limitations, and where the torque converter is at a disadvantage because of low power or low speeds, they will make headway. This type of performance seems outside the requirements of the construction machinery field, although possibly within the scope of agricultural tractor.

DISCUSSION

Questions Advantages Of Hydrostatic Transmissions

— H. W. Christenson
Allison Division, General Motors Corp.

(This discussion covers the Mortenson paper on pp 243–247, as well as the Bowers paper.)

I ADMIRE the accomplishments of the Dowty and Vickers companies in their application of hydrostatic drives to tractors and other working vehicles. However, as a manufacturer of torque converters, and full torque shifting transmissions, I wish to question some data and conclusions.

1. Efficiency.

In Fig. 13, Mr. Bowers shows the efficiency and drawbar pull curves on the hydrostatic unit blanketing the efficiency of several mechanical gears.

The hydraulic efficiency peaks at about 80%, whereas the mechanical gears peak at 70–75%. Our farm tractor mechanical gear trains exceed 75% overall efficiency. A 90% peak efficiency would be more reasonable for the mechanical gears and would be more consistent with his performance statement that the two tractors did comparable work while plowing a field.

2. Ratio coverage.

Ratio coverage of 4/1, or 5/1, is adequate for only a few vehicles. With torque converters, we have had to add 2–3–4 and sometimes 6 mechanical gear ratios giving a total coverage of about 20/1. To get this coverage with a hydrostatic motor would multiply the weight at least four or five times. We have several torque shifting transmissions which might be used to extend the hydrostatic ratio, but cost may prohibit hydrostatic competition to replace torque converters.

3. Control.

A lot has been said for the infinitely variable ratio control of hydrostatic transmissions, and I have to admire a drive that can give five complete reversals in 1 sec. However, on working machinery most drivers would have trouble killing or overspeeding their engines with positive ratio control and fast hydrostatic ratio changes can produce more shift shock than an engaging clutch. Whereas a torque converter adjusts ratio like a rubber band and automatically adjusts ratio thousands of times without waiting for an operator or engine speed governor to make a correction after an engine speed change. A ratio control lever would require more operator time and skill than our present forward reverse clutch lever.

Dynamic braking is claimed as an inherent feature of hydrostatic transmissions, but this feature requires a lot of safety controls to prevent overspeeding, overpressure, or overheating. All of which may be done, but not with the ease or capacity of a hydraulic retarder.

One feature of hydrostatic transmissions not mentioned in either paper is infinite speed ratio as the pump approaches zero displacement. This feature may be attractive to apply power with a gas turbine engine, whose rotor must idle from 20,000 to 40,000 rpm.

Mr. Bower's Closure To Discussion

1. Efficiency

The performance curves for the tractors fitted with the hydrostatic transmission and the mechanical transmission were both produced by the same independent tractor testing station. Some mechanical transmissions are not as efficient as we like to think, particularly if high viscosity oils are used at high gear speeds. This effect is not so pronounced with low viscosity hydrostatic transmission fluids.

2. Ratio Coverage

The hydrostatic transmission can provide ratios of 4 or 5/1 with a single variable unit, or up to 8/1 with two variable units, within the efficiency range shown. Vehicles with very wide ratio coverage often have two speed ranges, one for working and one for transport, and these can be provided by a simple two-speed shift on the hydraulic motor output. Ratio coverage in a mechanical transmission is sometimes extended to gain "creep speeds" which are automatically provided by the hydrostatic transmission. In those vehicles which spend the majority of their lives in a fixed top ratio at high speed, the efficiency under this condition is of greater importance than in the indirect ratios; and a torque converter with lockup provides a more suitable solution than a straight hydrostatic transmission.

3. Control

Because the hydrostatic transmission can hold a given ratio quite independently of load, rates of change of ratio can be controlled to any desired standard. If this type of performance is required, automatic ratio change to match load can easily be introduced by several simple methods, while still retaining a degree of driver control over ratio selection.

Vehicles are in production today which, except for emergency purposes, rely on the dynamic braking of the transmission. Providing that the vehicle weight and speed are low in relation to the power this is quite adequate. This applies mainly to such machines as bulldozers and farm tractors, where the normal load has virtually no positive inertia; though it is also true for road rollers because of their very low speed. Where the kinetic energy of the vehicle is high, the transmission braking must be supplemented by other means — if the braking energy cannot be absorbed by motoring the prime mover it must be dissipated as heat somewhere.

Hydrodynamic Split Torque Transmissions

Peter Bloch and Raymond C. Schneider
Twin Disc Clutch Co.

THIS PAPER deals with hydrodynamic split torque transmissions.

The first part is a mathematical investigation of systems with two parallel power paths of the power-dividing and power-uniting type. The basic equations for split torque transmissions as well as for engine-transmission packages are discussed. General criteria for the usefulness of split torque transmissions are presented.

The authors also investigate actual performance characteristics of various split torque transmissions. The overall effect of the engine transmission package on engine speed, output horsepower, output torque, and overall efficiency is analyzed and discussed.*

Mathematical Investigation

HYDRODYNAMIC SPLIT torque transmissions are power transmissions with two or more parallel paths of power flow and with a hydrodynamic coupling or a hydrodynamic torque converter built into one of these power paths. There is already a considerable amount of literature about hydrodynamic split torque transmissions, which are sometimes called differential or power-shunt transmissions.

This investigation is limited to the discussion of two path systems which consist of a hydrodynamic coupling or converter and a differential or planetary gearset that divides or unites the power flow.

Basic Equations

Figs. 1–6 show the split torque arrangements for which the following equations are applicable.

In Figs. 1–3, the planetary gearset has the function of a torque divider. One of the members of the planetary set is connected to the transmission input shaft, the second one drives the impeller of the hydrodynamic unit, and the third one is coupled to the transmission output shaft which is also connected to the turbine or the runner of the hydrodynamic unit. This arrangement shall be called "input split," and the ratio of impeller torque to input torque shall be designated a_1.

Fig. 1 shows true split torque arrangements, with $0 < a_1 < 0.5$ and $0.5 < a_1 < 1$. Part of the input power is going through the hydrodynamic unit, part of it is flowing through the mechanical power path.

Fig. 2 shows positive recirculative systems or regenerative systems as they are sometimes called, $1 < a_1 < 2$ and $2 < a_1 < \infty$. A certain part of the output power of the hydrodynamic unit returns to the planetary set. The power circulating through planetary gearset and hydrodynamic unit is, therefore, larger than the input power of the transmission. The energy flow through the hydrodynamic unit is positive, that is, from impeller to turbine.

Fig. 3 shows negative recirculative systems with $0 > a_1 > -1$ and $-1 > a_1 > -\infty$. A certain amount of the mechanically transmitted power is flowing in a negative direction through the hydrodynamic unit,

* Paper presented at SAE National Farm, Construction, and Industrial Machinery Meeting, Milwaukee, Sept. 16, 1959.

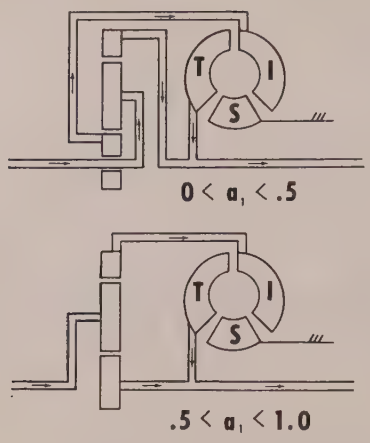


Fig. 1 — Input split torque arrangements — true split

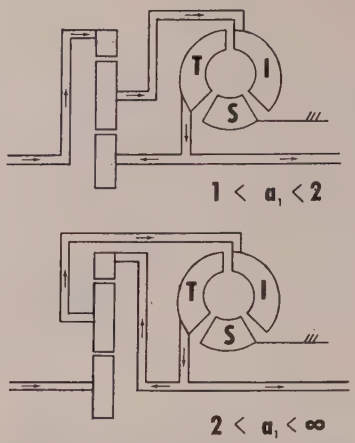


Fig. 2 — Input split torque arrangements — positive recirculative system

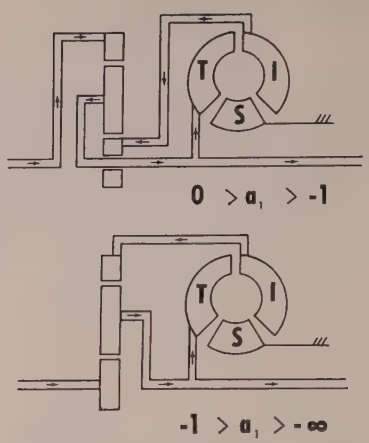


Fig 3 — Input split torque arrangements — negative recirculative system

from turbine to impeller and back to the planetary gearset.

The above classification of input splits is applicable to hydraulic couplings and to regular forward converters only. If a reverse converter is used, (that is, a converter whose impeller and turbine have an opposite sense of rotation), the negative recirculative system, for example, becomes either a true split or a positive recirculative system. The basic mathematical equations for that system, however, do not have to be modified.

In Figs. 4–6, the planetary gearset has the function of a torque collector. One of the members of the planetary set is connected to the transmission output shaft, the second one is driven by the turbine of the hydrodynamic unit, and the third is coupled to the transmission input shaft which also is connected to the impeller of the hydrodynamic unit. This arrangement shall be called "output split," and the ratio of turbine torque to output torque shall be designated a_2.

Fig. 4 again shows true split torque arrangements, with $0 < a_2 < 0.5$ and $0.5 < a_2 < 1$. Fig. 5 shows positive recirculative systems with $1 < a_2 < 2$ and $2 < a_2 < \infty$; and Fig. 6 represents negative recirculative systems with $0 > a_2 > -1$ and $-1 > a_2 > -\infty$.

This classification for output splits is applicable for hydraulic couplings and regular forward converters only. In order to establish the basic equations for split torque transmissions, we assume that the planetary gearset works at 100% efficiency and that all the losses of the transmission are concentrated in the hydrodynamic unit.

We need the following designations:

N = Speed
T = Torque
P = Power

dices:
1 = Transmission input
2 = Transmission output
p = Pump or impeller
t = Turbine or runner
m = Mechanical
con = Converter
tot = Transmission, total
max = Maximum
M = Match point
E = Engine rated

We define the following ratios:

$a_1 = \frac{T_p}{T_1}$ Input split ratio or ratio of impeller torque to input torque for input split.

$a_2 = \frac{T_t}{T_2}$ Output split ratio or ratio of turbine torque to output torque for output split.

$SR_{con} = \frac{N_t}{N_p}$ Converter speed ratio.

$SR_{tot} = \frac{N_2}{N_1}$ Transmission or total speed ratio.

$TR_{con} = \frac{T_t}{T_p}$ Converter torque ratio.

$TR_{tot} = \frac{T_2}{T_1}$ Transmission or total torque ratio.

$STR =$ Stall torque ratio (output shaft at stall).

$\eta_{con} = \frac{P_t}{P_p}$ Converter efficiency.

$\eta_{tot} = \frac{P_2}{P_1}$ Transmission or total efficiency.

Power Equations — The basic power equations are for input and output split:

Input power $P_1 = P_p + P_m$ (1)

Output power $P_2 = P_t + P_m$ (2)

Equations for Speed Ratios — We transform Eq. 1 into:

$$N_1 T_1 = N_p T_p + N_m T_m \quad (3)$$

With $a_1 = \frac{T_p}{T_1}$ (definition of input split),

$$1 - a_1 = \frac{T_1 - T_p}{T_1} = \frac{T_m}{T_1}, \text{ and}$$

$N_2 = N_m = N_t$ for input split, we obtain:

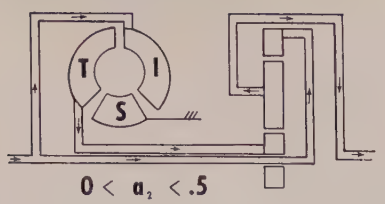
$$N_1 = N_p \frac{T_p}{T_1} + N_m \frac{T_m}{T_1} = N_p a_1 + N_2 (1 - a_1)$$

and:

$$\frac{N_1}{N_2} = \frac{N_p}{N_t} + a_1 \frac{N_2}{N_2} (1 - a_1) = \frac{N_p}{N_t} a_1 + 1 - a_1$$

$$\frac{1}{SR_{tot}} = \frac{a_1}{SR_{con}} + 1 - a_1$$

The relationship between converter speed ratio



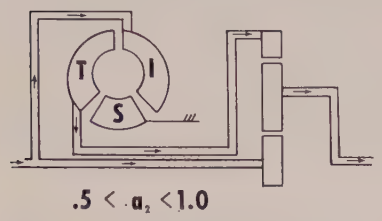
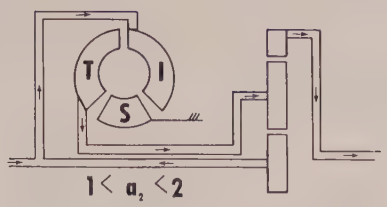


Fig. 4 — Output split torque arrangements — true split



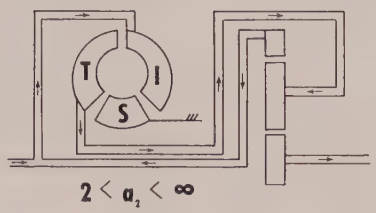
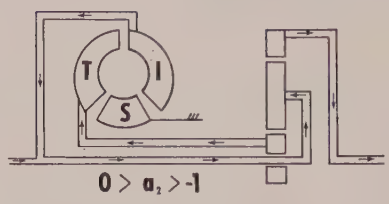


Fig. 5 — Output split torque arrangements — positive recirculative system



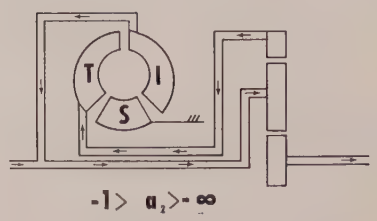


Fig. 6 — Output split torque arrangements — negative recirculative system

and transmission speed ratio for input split finally becomes:

$$SR_{tot} = \frac{SR_{con}}{a_1 + SR_{con}(1 - a_1)} \tag{4}$$

With

$$a_2 = \frac{T_t}{T_2} \text{ (definition of output split),}$$

$$1 - a_2 = \frac{T_2 - T_t}{T_2} = \frac{T_m}{T_2}, \text{ and}$$

$$N_1 = N_p = N_m \text{ for output split}$$

and transforming the Eq. 2 into:

$$N_2 T_2 = N_t T_t + N_m T_m \tag{5}$$

we obtain:

$$N_2 = N_t \frac{T_t}{T_2} + N_m \frac{T_m}{T_2} = N_t a_2 + N_1 (1 - a_2)$$

and:

$$\frac{N_2}{N_1} = \frac{N_t}{N_p} a_2 + 1 - a_2$$

The relationship between converter speed ratio and transmission speed ratio for output split finally becomes:

$$SR_{tot} = SR_{con} a_2 + 1 - a_2 \tag{6}$$

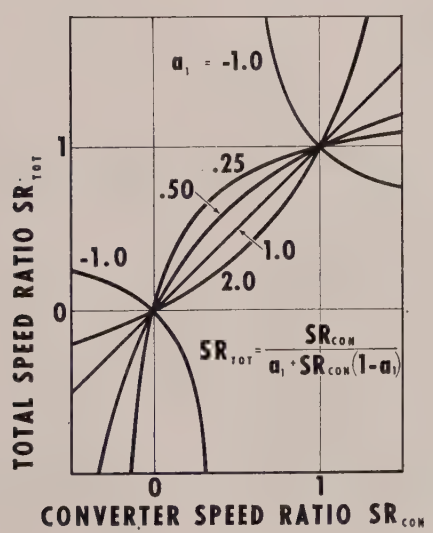


Fig. 7 — Speed ratio, input split

Eqs. 4 and 6 are plotted in Figs. 7 and 8 for different splits.

Fig. 7 shows that for the input split the total becomes zero if the converter speed ratio is zero and that for $SR_{con} = 1$ we get $SR_{tot} = 1$ as well. The first quadrant of Fig. 7 shows the normal operational field of a forward converter and a forward transmission; the second quadrant is the field of a reverse converter forming a forward transmission with negative split; the third quadrant contains the reverse converter producing a reverse transmission with positive split; and the fourth quadrant shows the field of a forward converter, making a reverse transmission with negative split. Fig. 7 shows also the points of infinity, in other words, the natural limitations of converter and transmission speed ratios.

Fig. 8 shows that for the output split, the converter speed ratio is not zero at stall of the package. However, for $SR_{con} = 1$ we find $SR_{tot} = 1$ for all values of split.

Equations for Torque Ratios — In the same manner and under the same conditions as above, we transform the Eq. 2 for the input split into:

$$T_2 = T_t + T_m \tag{7}$$

and:

$$\frac{T_2}{T_1} = \frac{T_t}{T_1} + \frac{T_1 - T_p}{T_1} = \frac{T_t T_p}{T_1 T_p} + 1 - \frac{T_p}{T_1}$$

The relationship between converter torque ratio

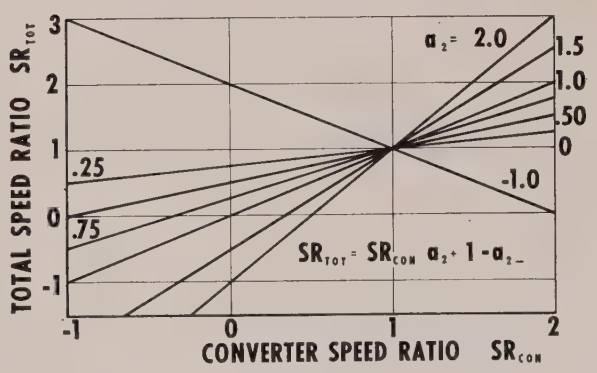


Fig. 8 — Speed ratio, output split

and total torque ratio becomes herewith:

$$TR_{tot} = TR_{con} a_1 = 1 - a_1 \qquad (8)$$

In a similar procedure, we transform Eq. 1 for the output split into:

$$T_1 = T_p + T_m \qquad (9)$$

and:

$$\frac{T_1}{T_2} = \frac{T_p}{T_2} + \frac{T_2 - T_t}{T_2} = \frac{T_p T_t}{T_2 T_t} + 1 - \frac{T_t}{T_2}$$

$$\frac{1}{TR_{tot}} = \frac{a_2}{TR_{con}} + 1 - a_2 = \frac{a_2 + TR_{con}(1 - a_2)}{TR_{con}}$$

The relationship between converter torque ratio and total torque ratio is finally:

$$TR_{tot} = \frac{TR_{con}}{a_2 + TR_{con}(1 - a_2)} \qquad (10)$$

In Figs. 9 and 10, Eqs. 8 and 10 are plotted with different splits as parameters. It may be pointed out that the basic structures of Eqs. 4 and 10, as well as 6 and 8 are identical.

It has to be underlined that for true input and output splits $0<a<1$ with forward converters, the torque multiplication of the transmission is always smaller than that of the converter alone, and that for positive splits $a>1$ the total torque ratio is larger than the converter torque ratio. For negative splits and reverse converters, the situation is more intricate.

Equations for Impeller Speed and Power for Input Split — Eq. 4, solved for SR_{con} becomes:

$$SR_{con} = \frac{SR_{tot} a_1}{1 - SR_{tot}(1 - a_1)} \qquad (11)$$

We divide this expression by SR_{tot} and obtain, since $N_2 = N_t$ for input split:

$$\frac{SR_{con}}{SR_{tot}} = \frac{N_t N_1}{N_p N_2} = \frac{N_1}{N_p} = \frac{a_1}{1 - SR_{tot}(1 - a_1)}$$

The impeller-input speed ratio is, therefore:

$$\frac{N_p}{N_1} = \frac{1 - SR_{tot}(1 - a_1)}{a_1} \qquad (12)$$

In order to obtain the impeller-input power ratio for input split, we have to multiply both sides of this equation with $\frac{T_p}{T_1} = a_1$:

$$\frac{P_p}{P_1} = 1 - SR_{tot}(1 - a_1) \qquad (13)$$

The impeller-input speed and power ratios are shown on Figs. 11 and 12. It might be pointed out that the impeller speed varies appreciably with total speed ratio and that the impeller power absorption is 100% at package stall, whereas it attains the magnitude of the split ratio at $SR_{tot} = 1$.

The corresponding equations for output split cannot be derived in a general manner, because the assumption of a specific converter characteristic becomes necessary.

Equations for Specific Converter

For the following mathematical investigations, the assumption of a specific converter characteristic is required.

In order to simplify the expressions, it is desirable to make use of a parabolic efficiency curve:

$$\eta_{con} = 4\eta_{con\,max} SR_{con}(1 - SR_{con}) \qquad (14)$$

Dividing this equation by SR_{con}, we obtain a linear function for the converter torque ratio:

$$\frac{\eta_{con}}{SR_{con}} = TR_{con} = 4\eta_{con\,max}(1 - SR_{con}) \qquad (15)$$

Besides this, we assume that our converters have a constant primary torque, that is, an input torque that does not vary with speed ratio.

The following equations are not accurately applicable for torque converters with different characteristics. They show, however, the general trends of split torque arrangements which are similar for all types of converters.

Impeller Power for Output Split — Since in output split arrangements the impeller rotates with input speed, the impeller-input power ratio is equal to the impeller-input torque ratio:

$$\frac{P_p}{P_1} = \frac{T_p}{T_1} \qquad (16)$$

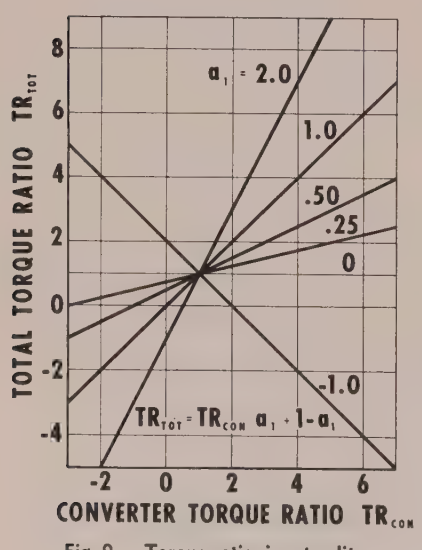


Fig. 9 — Torque ratio, input split

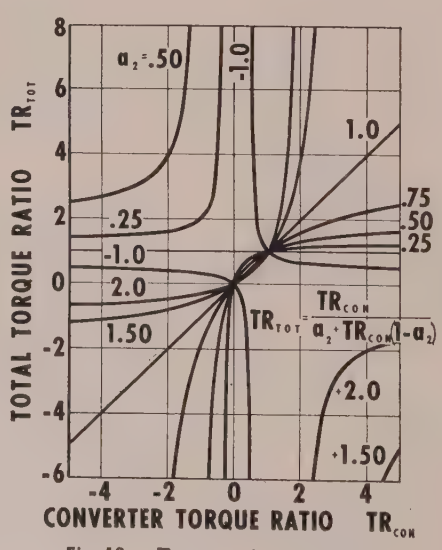


Fig. 10 — Torque ratio, output split

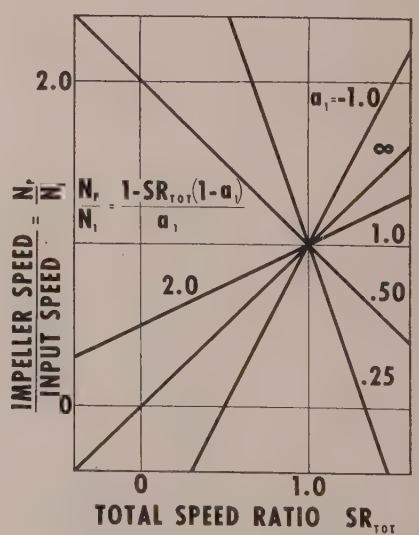


Fig. 11 — Impeller speed, input split

This expression leads to:

$$\frac{P_p}{P_1} = \frac{T_p}{T_1} \times \frac{T_t}{T_t} \times \frac{T_2}{T_2} = a_2 \frac{TR_{tot}}{TR_{con}} \tag{17}$$

With Eq. 10 for TR_{tot} we arrived at:

$$\frac{P_p}{P_1} = \frac{a_2}{a_2 + TR_{con}(1 - a_2)} \tag{18}$$

Introducing the specific converter characteristic as shown in Eq. 15 we obtain:

$$\frac{P_p}{P_1} = \frac{a_2}{a_2 + 4\eta_{con\,max}(1 - SR_{con})(1 - a_2)} \tag{19}$$

Transforming converter speed ratio into total speed ratio with Eq. 6, we write down the final result:

$$\frac{P_p}{P_1} = \frac{a_2^2}{4(_{con\,max}(1 - SR_{tot})(1 - a_2) + a_2^2)} \tag{20}$$

Fig. 13 illustrates that at $SR_{tot} = 1$, the whole input power is absorbed by the impeller. At $SR_{tot} = 0$, the output power is equal to zero. The difference between input power and impeller absorption has, therefore, to be furnished or dissipated by the converter turbine.

Stall Torque Ratio for Output Split — Since in output split arrangements the transmission stall point does not coincide with the converter stall point, the ratio between transmission stall torque and converter stall torque is not determined by Eq. 10:

$$TR_{tot} = \frac{TR_{con}}{a_2 + TR_{con}(1 - a_2)} \tag{10}$$

Introducing our converter characteristic (Eq. 15) we get:

$$TR_{tot} \quad \frac{4\eta_{con\,max}(1 - SR_{con})}{a_2 + 4\eta_{con\,max}(1 - SR_{con})(1 - a_2)} \tag{21}$$

According to Eq. 6 the converter speed ratio at transmission stall is:

$$SR_{con} = \frac{a_2 - 1}{a_2} \tag{22}$$

We substitute this expression in Eq. 21 and obtain the transmission stall torque ratio:

$$STR_{tot} = \frac{4\eta_{con\,max}}{a_2^2 + 4\eta_{con\,max}(1 - a_2)} \tag{23}$$

Since Eq. 15 determines the converter stall torque ratio:

$$STR_{con} = 4\eta_{con\,max}$$

we obtain the relationship between transmission and converter stall torque ratios as follows:

$$\frac{STR_{tot}}{STR_{con}} = \frac{1}{a_2^2 + 4\eta_{con\,max}(1 - a_2)} \tag{24}$$

For input split arrangements, this ratio is determined by Eq. 8 because here the turbine is connected to the transmission output shaft, and converter and transmission reach the stall point at the same time.

Fig. 14 shows the transmission-converter stall torque relationship for input and output splits, with a converter with 85% peak efficiency and a stall torque ratio of 3.4/1. It is interesting to see that the curve for output splits has a maximum at $a_2 = 1.7$.

Fig. 14 indicates that a real split means a loss of stall torque ratio and that a positive recirculative system allows a gain of stall torque ratio, for output splits at least in a certain practical useful field.

It has further to be pointed out that the loss of stall torque ratio in real splits is somewhat smaller for input split arrangements and that the gain of stall torque ratio in positive recirculative systems is more attractive in output split arrangements.

Efficiency — The total transmission efficiency is:

$$\eta_{tot} = \frac{P_2}{P_1} = \frac{N_2 T_2}{N_1 T_1} = SR_{tot} \times TR_{tot} \tag{25}$$

Using the Eq. 8 we obtain for the input split arrangement:

$$\eta_{tot} = SR_{tot}(TR_{con}\, a_1 + 1 - a_1)$$

$$= SR_{tot}\left(\frac{\eta_{con}}{SR_{con}} a_1 + 1 - a_1\right)$$

and with Eq. 11:

$$\eta_{tot} = SR_{tot}\left[\eta_{con} \frac{1 - SR_{tot}(1 - a_1)}{SR_{tot}} + 1 - a_1\right]$$

$$\eta_{tot} = \eta_{con}[1 - SR_{tot}(1 - a_1)] + SR_{tot}(1 - a_1) \tag{26}$$

This equation is applicable for any kind of converter.

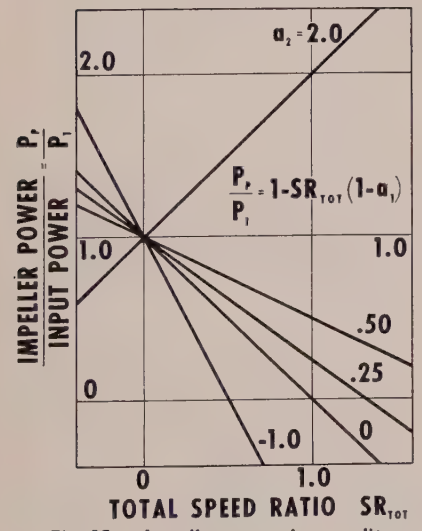


Fig. 12 — Impeller power, input split

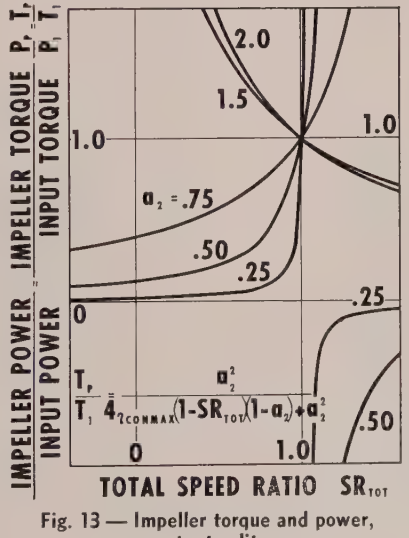


Fig. 13 — Impeller torque and power, output split

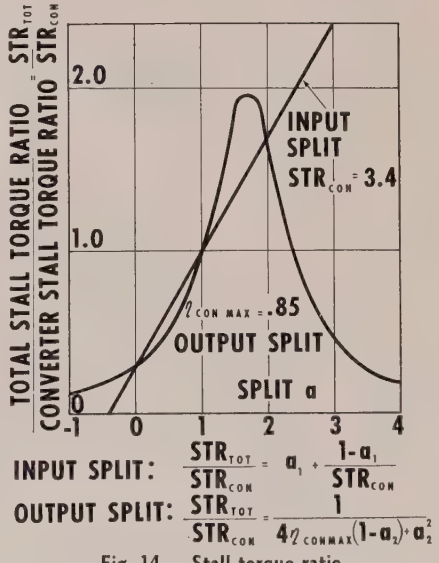


Fig. 14 — Stall torque ratio

Introducing Eq. 14 transformed with Eq. 11 we get for our specific case of converter:

$$\eta_{tot} = 4\eta_{con\,max}\frac{a_1 SR_{tot}(1-SR_{tot})}{1-SR_{tot}(1-a_1)} + SR_{tot}(1-a_1) \qquad (27)$$

For the output split arrangement we make use of Eq. 10 and write:

$$\eta_{tot} = SR_{tot}\frac{TR_{con}}{a_2 + TR_{con}(1-a_2)} = SR_{tot}\frac{\frac{\eta_{con}}{SR_{con}}}{a_2 + \frac{\eta_{con}}{SR_{con}}(1-a_2)}$$

$$= SR_{tot}\frac{\eta_{con}}{a_2 SR_{con} + \eta_{con}(1-a_2)}$$

We solve Eq. 6 for SR_{con}:

$$SR_{con} = \frac{SR_{tot} - 1 + a_2}{a_2} \qquad (28)$$

and obtain:

$$\eta_{tot} = \frac{\eta_{con} SR_{tot}}{\eta_{con}(1-a_2) + SR_{tot} - 1 + a_2} \qquad (29)$$

This equation is again applicable for any kind of converter. Using Eqs. 14 and 28, we obtain for our specific converter:

$$\eta_{tot} = \frac{SR_{tot}(1-SR_{tot})}{(1-SR_{tot})(1-a_2) + \frac{{a_2}^2}{4\eta_{con\,max}}} \qquad (30)$$

Eqs. 27 and 30 are illustrated in Figs. 15 and 16. Since in true split torque arrangements only a part of the input power is flowing through the converter

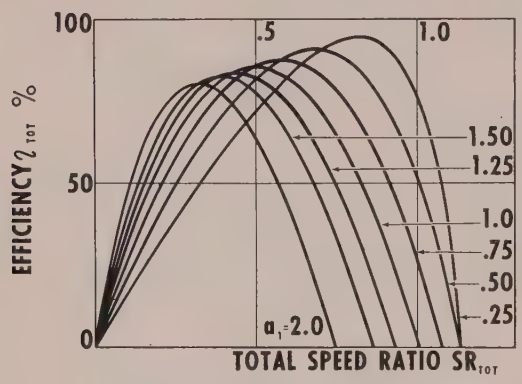


Fig. 15 — Efficiency input split

$$\eta_{tot} = \eta_{con\,max}\frac{a_1 SR_{tot}(1-SR_{tot})}{1-SR_{tot}(1-a_1)} + SR_{tot}(1-a_1)$$

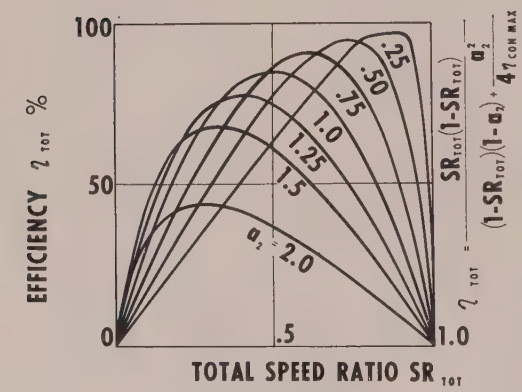


Fig. 16 — Efficiency output split

and since, therefore, only a part of this power is subject to losses, the peak transmission efficiency has to be higher than the peak converter efficiency. In recirculative systems, the converter power being higher than the transmission input power, the peak transmission efficiency is lower than that of the converter. It is interesting to see that the points of maximum efficiency are moving towards higher speed ratios with decreasing split ratios. It may also be pointed out that with decreasing split ratios, the broadness of the curves is diminishing.

Equations for Engine-Converter Packages

The following investigations are made for split torque arrangements which are connected to an internal-combustion engine. We assume that this engine has a constant output torque curve T_E = const. The engine rated power be P_E at the speed N_E and the torque T_E.

$$P_E = N_E T_E = N_E \times \text{const.} \qquad (31)$$

Engine Speed — Since the transmission input torque is constant, $T_1 = T_E$ = const, we can write for the input split arrangements:

$$T_p = a_1 T_1 = a_1 T_E = \text{const} \qquad (32)$$

If we call the transmission speed ratio, at which the transmission input power equals the engine rated power, the "match point" and if T_{pM} is the impeller torque at this match point, we obtain, since T_p = const:

$$T_{pM} = T_p \qquad (33)$$

Since the impeller torque depends for our specific converter on impeller speed only and not on converter speed ratio as well, as outlined in the last paragraph, we conclude:

$$N_{pM} = N_p \qquad (34)$$

In Eq. 12 we determined the ratio between impeller speed and input speed:

$$\frac{N_p}{N_1} = \frac{1-SR_{tot}(1-a_1)}{a_1} \qquad (12)$$

Transformed for the match point, this equation becomes:

$$\frac{N_{pM}}{N_E} = \frac{1-SR_{tot\,M}(1-a_1)}{a_1} \qquad (35)$$

Dividing Eq. 35 by Eq. 12 we get the ratio between actual engine speed and engine rated speed for any speed ratio for a given split a_1 and a given match point $SR_{tot\,M}$:

$$\frac{N_1}{N_E} = \frac{1-SR_{tot\,M}(1-a_1)}{1-SR_{tot}(1-a_1)} \qquad (36)$$

For practical purposes, it is sometimes more useful to give the match point in terms of converter speed ratio rather than transmission speed ratio. Making use of Eq. 4 we transform the Eq. 36, therefore, into:

$$\frac{N_1}{N_E} = \frac{1}{[1-SR_{tot}(1-a_1)] \times \left[1 + SR_{con\,M}\frac{1-a_1}{a_1}\right]} \qquad (37)$$

In a similar procedure, we write for the output split torque arrangements:

$$T_1 = T_p + T_m = T_E = \text{const} \qquad (38)$$

Since $T_m = T_t\frac{1-a_2}{a_2} = T_p TR_{con}\frac{1-a_2}{a_2}$

we arrive at:

$$T_1 = T_p \left[1 + TR_{con} \frac{1 - a_2}{a_2} \right] = T_E$$

and with Eq. 15 for our specific converter we obtain:

$$T_1 = T_p \left[1 + 4\eta_{con\,max} \frac{1 - a_2}{a_2} (1 - SR_{con}) \right] = T_E \quad (39)$$

or, transformed with Eq. 6:

$$T_1 = T_p \left[1 + 4\eta_{con\,max} \frac{1 - a_2}{a_2^2} (1 - SR_{tot}) \right] = T_E \quad (40)$$

Eq. 39, written for the converter match point, becomes:

$$T_E = T_{pM} \left[1 + 4\eta_{con\,max} \frac{1 - a_2}{a_2} (1 - SR_{con\,M}) \right] \quad (41)$$

Since impeller speed is identical with input speed, we have, according to the basic relationship between speed and torque in a hydrokinetic machine:

$$\frac{T_p}{T_{pM}} = \left(\frac{N_1}{N_E} \right)^2 \quad (42)$$

By calculating the ratio T_p/T_{pM} with Eqs. 40 and 41, we obtain the final result:

$$\frac{N_1}{N_E} = \sqrt{\frac{1 + 4\eta_{con\,max} \frac{1 - a_2}{a_2} (1 - SR_{con\,M})}{1 + 4\eta_{con\,max} \frac{1 - a_2}{a_2^2} (1 - SR_{tot})}} \quad (43)$$

The ratios between actual engine speed and rated engine speed for input and output splits are shown in Figs. 17 and 18 for two different amounts of split and for two different match points. We can see that for true splits, the engine speed is going up with increasing total speed ratio until it reaches the rated speed. After that, the engine governor keeps the speed practically constant. In this phase, of course, the transmission power and torque are rapidly decreasing. For positive recirculative systems, the engine speed is decreasing speed ratio. At speed ratios below the match point, the engine is running on the governor with reduced torque and power.

Converter Size — The basic relationship between torque, speed and diameter of a hydrokinetic machine is:

$$T = kN^2D^5 \quad (44)$$

k is a design constant which is characterized by the converter blading, the converter channel area, the converter circuit shape, and things of that nature.

If we are going to investigate the influence of split torque arrangements on converter size, we have, of course, to work with converters with the same specific characteristics, which means that we are not allowed to change the design constant k.

We write Eq. 44 for a converter in a split torque arrangement at the match point of the transmission:

$$T_{pM} = kN_{pM}^2D_p^5 \quad (45)$$

and for a converter with direct connection to an identical prime mover:

$$T_1 = kN_1^2D_1^5 \quad (46)$$

Since we are dealing with a converter with constant input torque and an engine with constant output torque, the match point is immaterial for Eq. 46.

Dividing Eq. 45 by Eq. 46 and solving for the size ratio D_p/D_1, we obtain:

$$\frac{D_p}{D_1} = \sqrt[5]{\frac{T_{pM}N_1^2}{T_1N_{pM}^2}} \quad (47)$$

For input split arrangements, we have, since the input torque of the split torque drive is equal to that of the straight converter:

$$\frac{T_{pM}}{T_1} = a_1$$

$\frac{N_{pM}}{N_1}$ is given by Eq. 12.

The size ratio becomes, therefore:

$$\frac{D_p}{D_1} = \sqrt[5]{\frac{a_1^3}{[1 - SR_{tot\,M}(1 - a_1)]}} \quad (48)$$

Transformed with Eq. 4 we obtain:

$$\frac{D_p}{D_1} = \sqrt[5]{a_1[SR_{con\,M}(1 - a_1) + a_1]^2} \quad (49)$$

For output split arrangements the impeller speed is equal to the engine speed, and since the input torque of the split torque drive is again equal to that of the straight converter drive, the ratio T_{pM}/T_1 is determined by Eqs. 16 and 19.

We obtain, therefore:

$$\frac{D_p}{D_1} = \sqrt[5]{\frac{1}{1 + 4\eta_{con\,max} \frac{1 - a_2}{a_2} (1 - SR_{con\,M})}} \quad (50)$$

The Eqs. 49 and 50 are presented in Figs. 19 and 20. For a straight converter drive, split $a = 1$, the size

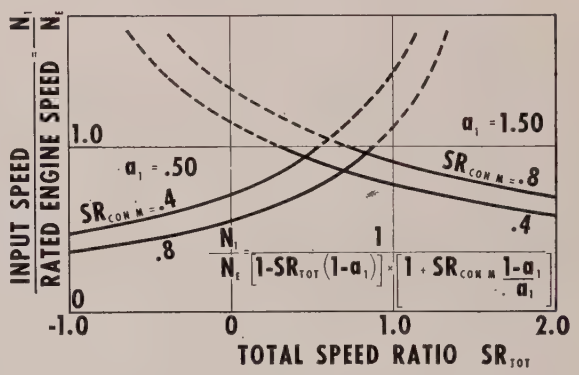


Fig. 17 — Input speed, input split

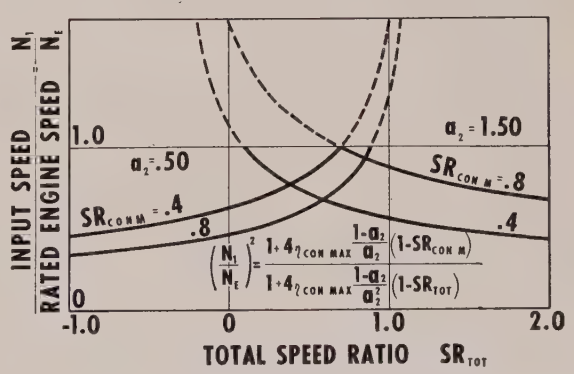


Fig. 18 — Input speed, output split

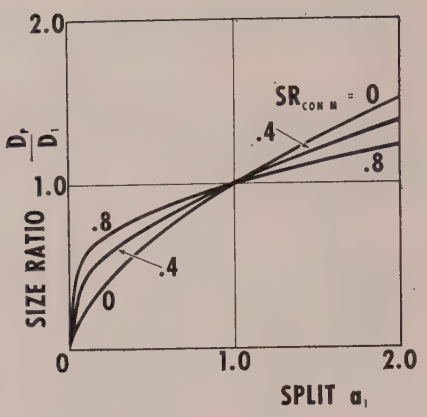
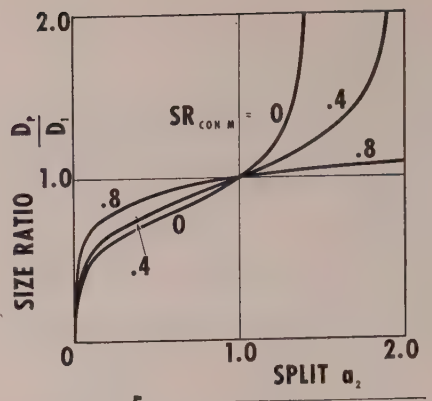


Fig. 19 — Converter size, input split

$$\frac{D_p}{D_1} = \sqrt[5]{a_1\left[SR_{CON\ M}(1-a_1) + a_1\right]^2}$$



$$\frac{D_p}{D_1} = \sqrt[5]{\frac{1}{1 + 4\eta_{CON\ MAX}\frac{1-a_2}{a_2}(1-SR_{CON\ M})}}$$

Fig. 20 — Converter size, output split

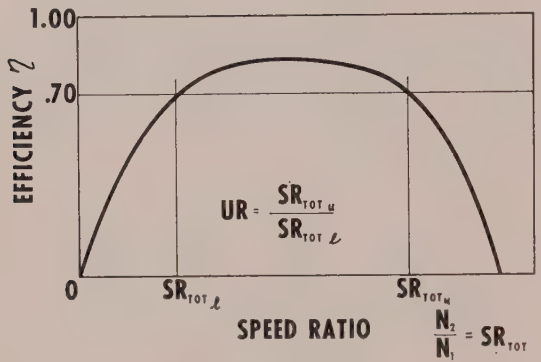


Fig. 21 — Utility ratio

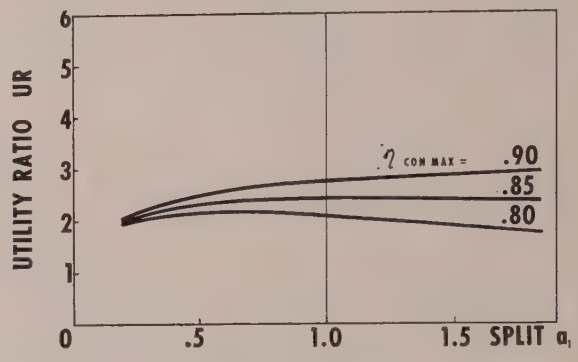


Fig. 22 — Utility ratio, input split

ratio is of course $D_p/D_1 = 1$.

In the field of real splits $a < 1$, the size ratio is decreasing with decreasing split ratio, the curves with lower match points showing a lower size ratio. In the field of positive recirculation, the size ratio increases with increasing split ratio, the curves with lower match point showing a faster increase of size ratio. The points of infinity which occur in output split arrangements do correspond to the points of infinity in Fig. 13.

It may be pointed out that a reduction of converter size is easier to attain with input split arrangements. The size ratio for a split $a_1 = 0.25$ and a converter matched at $SR_{con\,M} = 0.4$ is 60%, which represents a reduction of converter diameter of 40%.

General Criteria for Usefulness of Split Torque Transmissions

The main purpose of the installation of a hydrokinetic torque converter is to adjust the output speed automatically to the output load transmitting as much of the input power as possible. Emphasis may be given either to an efficiency curve as high as possible over a range as broad as possible or to the ability of the transmission to put out as much power as possible over a range as large as possible. It will be shown that these two features are not necessarily identical.

In order to discuss the influence of split torque arrangements upon efficiency and output horsepower characteristics, we introduce the terms of utility ratio, economy ratio, and availability ratio.

Utility Ratio — Fig. 21 shows the efficiency curve of a torque converter or split torque arrangement plotted versus speed ratio. We arbitrarily decide that the practically useful range of this transmission be the speed range between the two 70% points of the efficiency curve.

If $SR_{tot\,u}$ and $SR_{tot\,l}$ are the speed ratios at the upper and at the lower 70% points, we define the utility ratio to be:

$$UR = \frac{SR_{tot\,u}}{SR_{tot\,l}} \tag{51}$$

Solving the efficiency equation for input splits (Eq. 27) for SR_{tot} with $\eta_{tot} = 0.70$, we obtain for the upper and the lower 70% points:

$$SR_{tot\,u_l} = \frac{1 \pm \sqrt{1 - \dfrac{2.8[a_1^2 + a_1(4\eta_{con\,max} - 2) + 1]}{[a_1(4\eta_{con\,max} - 1.7) + 1.7]^2}}}{\dfrac{2[a_1^2 + a_1(4\eta_{con\,max} - 2) + 1]}{a_1(4\eta_{con\,max} - 1.7) + 1.7}} \tag{52}$$

The utility ratio for input split arrangements is, therefore:

$$UR = \frac{1 + \sqrt{1 - \dfrac{2.8[a_1^2 + a_1(4\eta_{con\,max}{}^{-2}) + 1]}{[a_1(4\eta_{con\,max}\,1.7) + 1.7]^2}}}{1 - \sqrt{1 - \dfrac{2.8[a_1^2 + a_1(4\eta_{con\,max}{}^{-2}) + 1]}{[a_1(4\eta_{con\,max}\,1.7) + 1.7]^2}}} \tag{53}$$

Making use of Eq. 30 we obtain in a similar pro-

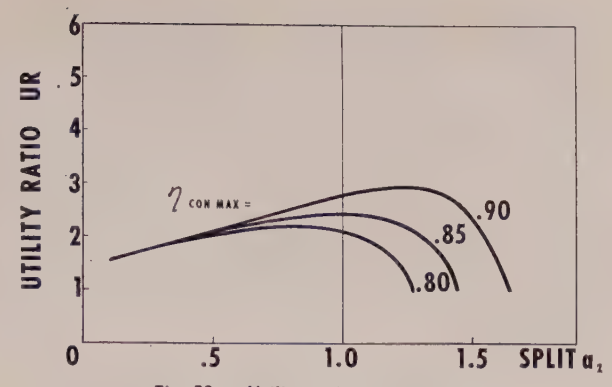


Fig. 23 — Utility ratio, output split

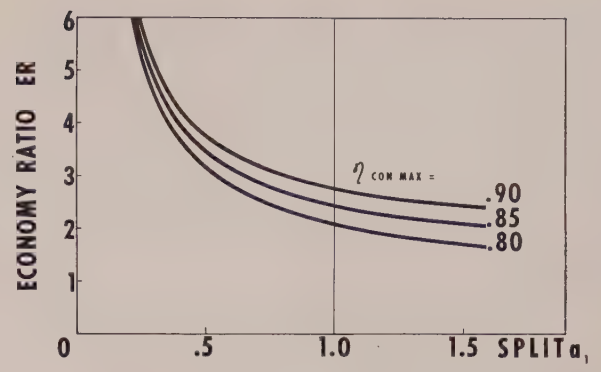


Fig. 25 — Economy ratio, input split

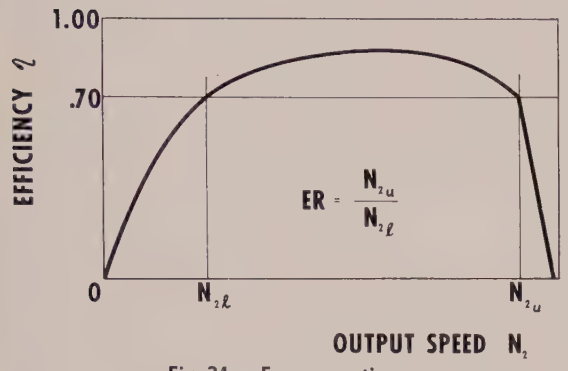


Fig. 24 — Economy ratio

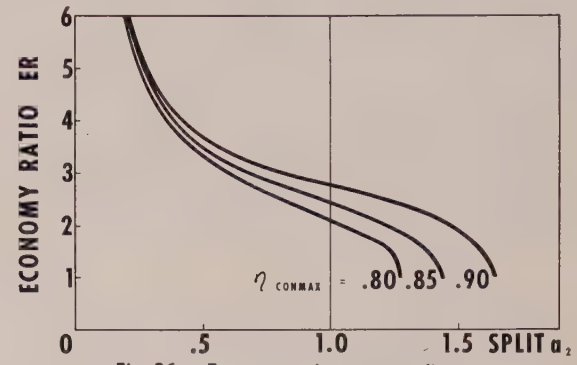


Fig. 26 — Economy ratio, output split

cedure for the output splits:

$$SR_{\text{tot}\,ul} = \frac{1 \pm \sqrt{1 - \dfrac{2.8(1-a_2)\eta_{\text{con max}} + 0.7a_2^2}{\eta_{\text{con max}}(1.7-0.7a_2)^2}}}{\dfrac{1.7-0.7a_2}{2}} \tag{54}$$

The utility ratio for output split arrangements becomes herewith:

$$UR = \frac{1 + \sqrt{1 - \dfrac{2.8(1-a_2)\eta_{\text{con max}} + 0.7a_2^2}{\eta_{\text{con max}}(1.7-0.7a_2)^2}}}{1 - \sqrt{1 - \dfrac{2.8(1-a_2)\eta_{\text{con max}} + 0.7a_2^2}{\eta_{\text{con max}}(1.7-0.7a_2)^2}}} \tag{55}$$

The Eqs. 53 and 55 are shown on Figs. 22 and 23 for several values of converter peak efficiency.

It is evident that a substantial gain of utility ratio cannot be obtained with split torque arrangements. In true splits only converters with low peak efficiency, and in positive recirculative systems only converters with extremely high peak efficiency will show some modest improvement. As a rule, it may be stated that in general, split torque arrangements are tending to reduce the utility ratio.

Economy Ratio — Fig. 24 shows the efficiency curve of a torque converter or split torque arrangement plotted versus output speed.

The economy ratio is defined as the ratio of the output speeds at the upper and at the lower 70% efficiency points. The values of output speed that correspond to the two 70% efficiency points are N_{2u} and N_{2l}.

$$ER = \frac{N_{2u}}{N_{2l}} \tag{56}$$

Under the condition that at the 70% efficiency points the engine speed is not limited by the engine governor, we obtain for input split arrangements, since:

$$\frac{N_2}{N_E} = \frac{N_2}{N_1} \times \frac{N_1}{N_E}$$

and with Eq. 37:

$$\frac{N_2}{N_E} = \frac{SR_{\text{tot}}}{[1 - SR_{\text{tot}}(1-a_1)] \times \left[1 + SR_{\text{con}\,M}\,\dfrac{1-a_1}{a_1}\right]} \tag{57}$$

With this expression written for the two 70% efficiency points, we arrive at the equation for the economy ratio for input split arrangements:

$$ER = \frac{SR_{\text{tot}\,u}}{SR_{\text{tot}\,l}} \times \frac{[1 - SR_{\text{tot}\,l}(1-a_1)]}{[1 - SR_{\text{tot}\,u}(1-a_1)]} \tag{58}$$

$SR_{\text{tot}\,u}$ and $SR_{\text{tot}\,l}$ are determined by Eq. 52.

In a similar manner, we obtain for output split arrangements, using Eq. 43

$$\frac{N_2}{N_E} = SR_{\text{tot}} \sqrt[2]{\frac{1 + 4\eta_{\text{con max}}\,\dfrac{1-a_2}{a_2}\,(1 - SR_{\text{con}\,M})}{1 + 4\eta_{\text{con max}}\,\dfrac{1-a_2}{a_2^2}\,(1 - SR_{\text{tot}})}} \tag{59}$$

and:

$$ER = \frac{SR_{\text{tot}\,u}}{SR_{\text{tot}\,l}} \sqrt[2]{\frac{1 + 4\eta_{\text{con max}}\,\dfrac{1-a_2}{a_2^2}\,(1 - SR_{\text{tot}\,c})}{1 + 4\eta_{\text{con max}}\,\dfrac{1-a_2}{a_2^2}\,(1 - SR_{\text{tot}\,u})}} \tag{60}$$

$SR_{\text{tot}\,u}$ and $SR_{\text{tot}\,l}$ being given by Eq. 54.

Eqs. 58 and 60 are illustrated in Figs. 25 and 26.

The rise of the economy ratio in true split torque arrangements with decreasing amount of split is remarkable. This phenomenon is not too difficult to understand; as shown in Figs. 17 and 18, the engine speed is severely lugged down at small speed ratios for splits $a < 1$. Hence, to arrive at a certain output speed, the converter runs at a higher speed ratio and herewith at a higher efficiency. The whole left end of the efficiency curve is, therefore, moved towards lower output speeds which increases our economy ratio.

In positive recirculative systems, the economy ratio is dropping, mainly because these systems are reducing the transmission peak efficiency.

The increasing economy ratio in true split arrangements is of foremost importance for transmissions where engine lug-down and, therefore, loss of input power at low speeds can be accepted and where high efficiency over a wide speed range is desired.

Availability Ratio — Fig. 27 shows the output power curve of a torque converter or split torque arrangement plotted versus output speed.

The availability ratio is defined as the ratio of the upper to the lower output speeds at which the output power equals 70% of the engine rated power.

$$AR = \frac{N_{2u}}{N_{2l}} \qquad (61)$$

In order to obtain the availability ratio as a function of the magnitude of output or input split, we have to calculate the output power curves of a given converter with a given match point for different splits. For each curve, we have to establish the availability ratio separately.

This has been done for a converter with parabolic efficiency curve and a peak efficiency of $\eta_{con\,max}$ = 0.90, 0.85, and 0.80, with a match point of $SR_{con\,M}$ = 0.80 and with regard to the influence of the engine governor.

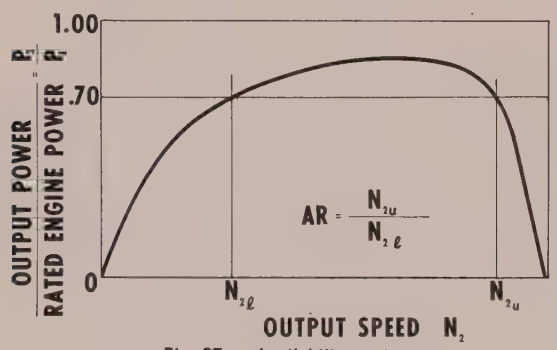


Fig. 27 — Availability ratio

The results are summarized in Figs. 28 and 29 for input and output split arrangements. They make it clear that the availability ratio is considerably reduced in all kinds of split torque arrangements, due to the reduction of engine power because of engine lug-down and engine governor action. Hence, for applications where a maximum possible of output power is desired at any output speed, as for instance in locomotives, split torque arrangements are not recommended.

Actual Performance Investigation

The actual application of a split torque transmission to an internal-combustion engine and the resulting performance characteristics will be investigated in the following pages:

Criteria for Configuration Selection

The inclusion of both input and output splits and the range of a passing from $-\infty$ to $+\infty$ gives an infinite number of possible mechanical and converter configurations. The actual choice of design configurations to be considered must be made chiefly on the basis of performance. Therefore, the first step in a general solution must be for the user to set forth the criteria or envelope of operating conditions required. These conditions could be lengthy and involved or could be quite simple. Oversimplification of operating criteria by the user is poor in that no specific path is pointed out for the designer to investigate and the resulting search for a solution is aimless and wasteful.

Some of the more pertinent operating data usually given before a design study can be undertaken are:

1. What output torque is necessary at stall in the vehicle and thru what speed range must the vehicle operate?
2. Is engine lug desirable or undesirable?
3. In what range will the vehicle operate most of the time?
4. Is the vehicle variable throttle operated or normally full throttle operated?
5. Is downhill braking important?
6. How important is availability ratio versus economy ratio?

With some or all of these requisites, the designer, with the aid of general equations and curves as shown in Figs. 1–29, can eliminate numerable variations that do not fulfill all the necessary require-

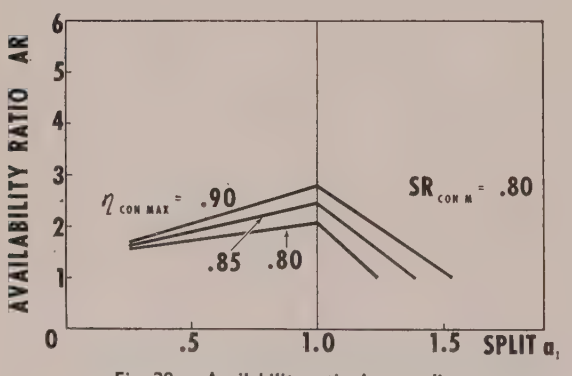


Fig. 28 — Availability ratio, input split

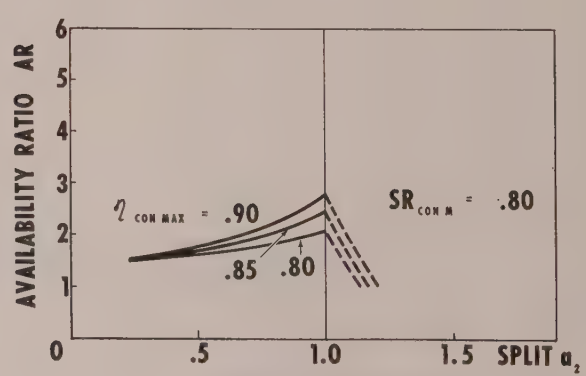


Fig. 29 — Availability ratio, output split

ments. The design investigation can then proceed with calculations of actual performance characteristics of the possible choices. This performance analysis, together with examinations by layout of the mechanical problems involved, will narrow down the choices and point the way to a final design.

True Input Splits

To examine the performance characteristics in the light of actual installations of engine transmission packages, we will first examine some "true" input splits where a_1 varies from 0 to + 1.

An input split where $a_1 = 0.70$ is pictured schematically in Fig. 30, the mechanical arrangement here is straightforward. The planetary gearset is between engine and converter and is sized to provide a 70% hydraulic-30% mechanical torque split. The torque flow through the unit is shown in shaded lines in the bottom half of the figure. The particular point described in the figure is at 0.6 converter speed ratio and a converter efficiency of 83%. The input torque to the converter is 70% of the engine torque and the converter then multiplies this torque by its torque ratio at the speed ratio given.

The total output torque is then a sum of the mechanically delivered torque and the converter turbine torque. The upper half of this figure shows the power flows through the transmission at the same instant. The dashed lines indicate the power paths through the unit at the point under examination. It will be noted that in addition to the 83% efficiency of the torque converter, the gearset is credited with a 98% efficiency. Therefore, 2% of the engine power is considered lost in gearing inefficiency. The actual gearing efficiency may vary considerably from this figure as any clutch drags, piston-ring seals, oil distributors, viscous drags, and the like, have losses which must be considered in the overall package. The mechanical path is shown transmitting some 20% of the power and the hydraulic path some 78%. The overall efficiency of this package is then 85% versus 83% efficiency for the converter alone. This figure compares well with the efficiency changes shown in Fig. 15, considering a loss for the gearing. A power analysis of the input split unit shown in Fig. 30 is plotted for the complete range of package speed ratios in Fig. 31. This curve shows how the mechanical per cent of the power varies linearly with speed ratio and is an actual application plot of a specific split shown generally in Fig. 12.

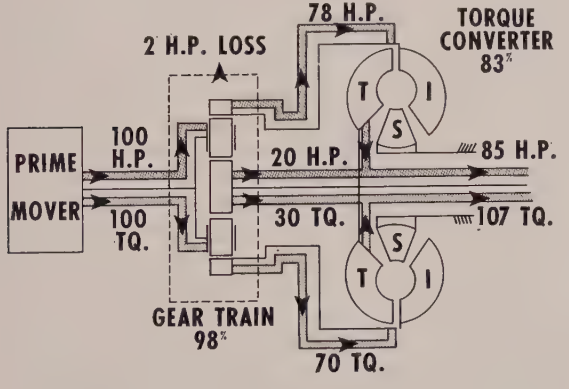


Fig. 30 — Input split — power and torque flow ($a_1 = 0.70$)

The input split package pictured schematically in Fig. 32A is quite similar to that shown in Fig. 30. The performance characteristics have been completely shown in Fig. 32B and for this reason an actual converter is shown with a freewheel between the impeller and the turbine. The converter shown is a single-stage stationary housing type with a flat primary curve and a sustained torque characteristic at zero efficiency and high speed ratio.

The freewheel in this installation then serves two purposes:

1. It acts as a positive mechanical push-to-start device between the fluid drive and the engine.
2. It serves to lock impeller and turbine together at 1.0 converter speed ratio and, therefore, keeps the package at 1.0 speed ratio and at the efficiency occurring there. This arrangement is useful for high-speed, low-load conditions or for full-throttle, no-load condition where all the power put into the unit turns into heat. If the freewheel were not included, the output shaft speed (and total speed ratio) would rise at no load until the turbine absorption was equal to the mechanical path input torque, and the amount of power dissipated to heat would rise substantially.

The location of the freewheel in this arrangement is arbitrary, as it could be between the sun gear and

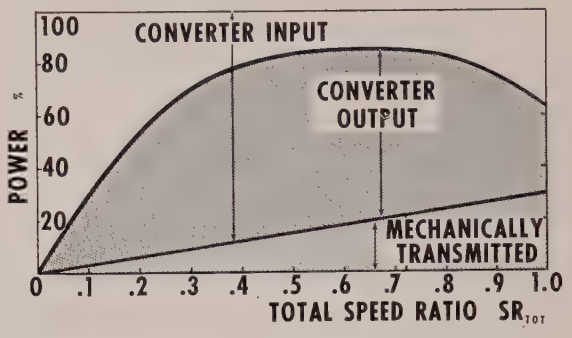


Fig. 31 — Input split — power analysis ($a_1 = 0.70$)

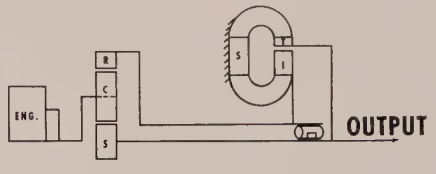


Fig. 32A — Input split schematic ($a_1 = 0.66$)

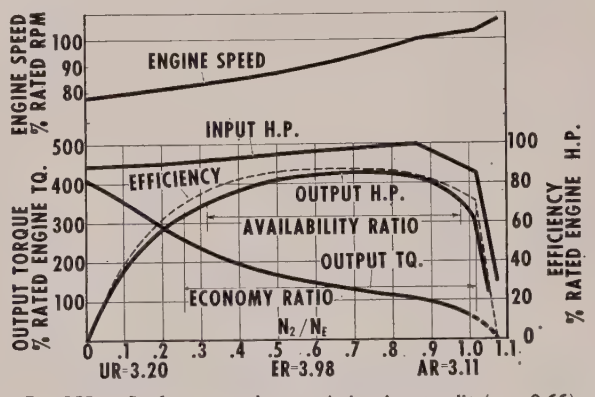


Fig. 32B — Performance characteristics, input split ($a_1 = 0.66$)

carrier or even between the ring gear and the carrier. It will be noted, however, that at different locations the freewheel absorbs varying percentages of full torque. The location shown allows the freewheel to transmit only 66% of the total torque carried back to the engine. This location is important for relative freewheel size and strength.

Without some sort of protective device between the freewheel and the engine, the freewheel would be forced to take the total instantaneous transmitted torque of any clutch between it and the ground. With power-shift transmissions having large, wet, multiple plate clutches in the following transmission or final drive, these torques could easily be over the designed capacity of the freewheel. A simple slip clutch or some type of degenerative clutch could be used between the planetary and the engine to protect the freewheel in installations of this sort.

The performance curve Fig. 32B is typical for a "true" input split $a_1 = 0.66$ when matched to a normal diesel engine curve and using a stationary housing single-stage, flat primary curve converter. The particular converter used in this illustration had a peak efficiency of 84.5% and a stall torque ratio of 5.0/1.

The engine speed has been lugged down some 20% at stall. This lug naturally causes a decrease in engine power shown on the input power curve. If the engine had an exceedingly high torque rise so that the horsepower curve was flat for speeds less than rated, the lug would not necessarily decrease input power, and the resultant output power curve would be higher at the lower output speeds. This type of engine curve would automatically increase availability ratio. The economy ratio previously defined in Fig. 24 is shown at the bottom of Fig. 32B. The availability ratio as defined in Fig. 27 is shown at the center of the curve. The question of how important each of these is to the user will now be reviewed. The availability ratio is a measure or index of the amount of usefulness in output power of the *installed engine power* in the vehicle. For most industrial installations that operate continuously through a wide band of output speeds at full throttle, the availability ratio should be a very meaningful and useful index. The economy ratio is a comparative measure of the width of the efficiency curve of an engine-transmission package. For installations where fuel economy and transmission efficiency are of foremost importance, this should be a useful tool. The input horsepower and engine speed curves at the high output speeds show how the usual governor characteristic and the freewheel affect the high end of the curve. The input power curve shows that the engine reaches full governed speed and horsepower at 0.86 output speed. The next portion of the power line is a result of the normal governor characteristics being applied to the transmission package. At 1.015 output speed the freewheel locks the package at 1.0 speed ratio and the engine continues down the governor curve until zero output power in the package is achieved. Since the converter used has a rather flat primary torque curve, there is some absorption at 1.0 speed ratio. With zero output power the input power curve shows approximately 30% of the engine rated power still being absorbed by the converter. The type of converter used and the amount of split can vary this figure of engine power dissipated to heat all the way from zero to 100%.

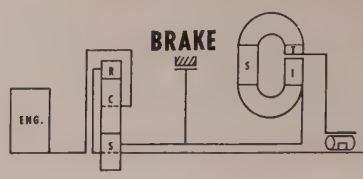


Fig. 33A — Input split schematic ($a_1 = 0.25$)

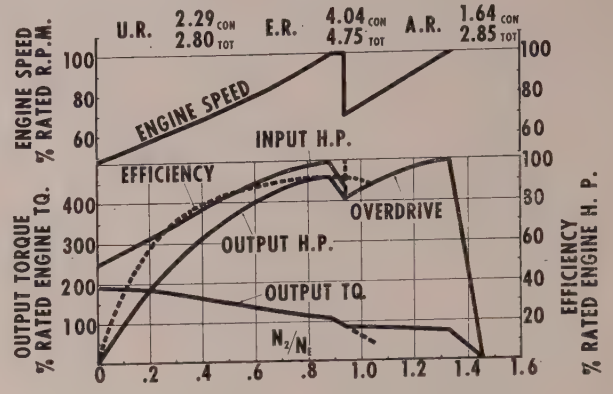


Fig 33B — Performance characteristics, input split with overdrive ($a_1 = 0.25$)

Its value and its effect on the overall installation is an important design point.

Another illustration of a "true" input split with $a_1 = 0.25$ is shown schematically in Fig. 33A. Since the requirement of hydraulic torque is only 25% of input torque, the planetary is connected in a manner opposite that shown in Fig. 32A.

An overdrive may then be obtained with the configuration simply by locking the sun gear which gives an overdrive ratio through the ring gear (output shaft) of $\frac{1}{\%\ \text{Mechanical}}$ or $\frac{1}{0.75} = 1.33/1$. Thus, the sun gear is locked by the brake shown and to prevent hydraulic losses (due to the turbine rotating in the converter at this time) a freewheel is placed between the turbine and the output shaft to allow the turbine to stop when the output shaft overspeeds it. The mechanical overdrive range then adds considerably to the overall performance at very small additional expense or complication.

The freewheel in this installation is not at all in the same severe application as was the freewheel in Fig. 32A. In this application the freewheel must be designed only to transmit full turbine torque and is not ever subjected to high inertia shock loading.

It may be pointed out that this basic configuration may be varied in many ways to obtain performance variations. A clutch could be added between any two members of the planetary to give a direct drive feature. This addition however, will probably lead to further additional clutches or dumping converters to drop the hydraulic loss in direct drive.

The "Diwabus" transmission manufactured by J. M. Voith G.M.B.H., Heidenheim (Brenz), Germany, is a split torque drive quite similar to that pictured. The Voith unit is used for buses and small railcars. It has a split of $a_1 = 0.50$ and attains this split with a planetary using a stepped pinion and two sun gears. This makes the overdrive ratio in this unit 2.0/1.

A performance curve for the low hydraulic input split is shown in Fig. 33B. This again uses a sta-

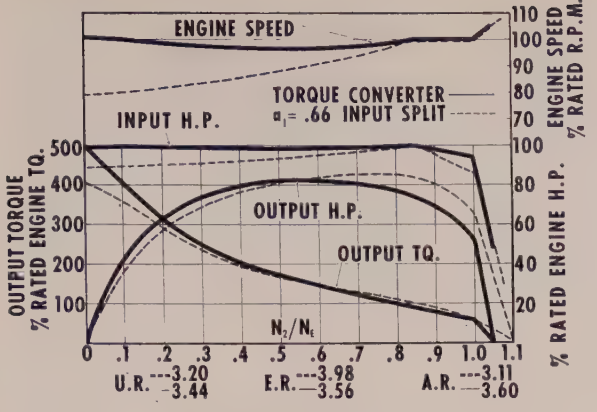


Fig. 34 — Performance comparison, input split versus torque converter

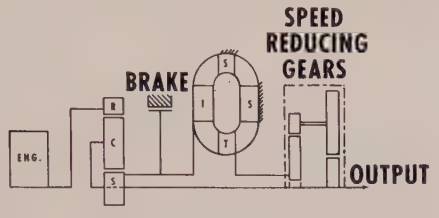


Fig. 35A — Input split, Ilomatic transmission ($a_1 = -0.57$)

tionary housing single-stage converter with a flat primary torque curve. The severe engine lug shown in the hydraulic range results in a considerable loss in engine input power. The curve shows the overdrive range and the resulting "hole" in the output horsepower curve at the shift point. In an effort to keep the transmission in the high efficiency range continuously, designers have changed converter match points or added direct drives to this type of split. The addition of a direct drive fills the "hole" in the output curve fairly effectively and also allows the converter to be matched at a lower speed ratio. This gives less severe engine lug, a smaller converter, and a greater availability ratio. To pay for these performance increases, more mechanical complication is required in the form of one or two more clutches, and the resulting mechanical drag of these clutches when they are disengaged in the other phases of operation. The output torque at stall is never very high due to the small percentage of torque being multiplied by the converter.

A problem on any split torque unit that is combined with overdrive or direct drive phases is the control of the shift points. The control systems to automatically upshift and downshift this type package from hydraulic to direct to overdrive and back are quite complicated. The cost and added complication of this system must be considered when comparing various types of hydraulic transmissions.

The comparison of an input split $a_1 = 0.66$ with a pure torque converter application is made in Fig. 34. The converter used in both the input split torque and the pure converter application is a stationary housing single-stage torque converter. The result of the flat primary curve of this type unit is graphically shown in the engine speed and input power curves. The effect of the split on engine speed and input power is easily seen. The usual lowering of output torque at stall is present as is the shift of the output horsepower to the right.

This input split arrangement has made the converter size some 10% smaller, as was shown generally in Fig. 19. The heat dissipation mentioned for Fig. 32B is shown as a direct comparison here. The pure converter installation at zero output power still absorbs 47% of the rated engine power while the split torque unit lowers this value to 30%.

The comparisons of the utility, economy, and availability ratios are made at the bottom of the figure. These show the usual relationship that is generally true for input and output splits with $0 < a < 1$. *A decrease in utility ratio and availability ratio* and an increase in economy ratio is typical of these splits. The use of a converter with a flat primary torque curve is assumed in this generalization. The use of a converter curve with absorption characteristics that drop from stall to high speed ratios (such as typical multistage units with a turbine stage directly preceding the impeller) will also make all these relationships true for the areas $0 < a < 1$.

Converters with rising absorption characteristics from stall to high speed ratios will *not* follow this generalization on utility, economy, and availability ratios. These converters are usually built specifically for split torque applications and would not usually have many applications as pure converters.

Recirculative Input Splits

To show an actual example of one of the negative input splits, the "Ilomatic" transmission manufactured in Pinneberg, Germany, by Ilowerke will be discussed. Fig. 35A shows a schematic diagram of this transmission. This arrangement is similar to the negative recirculative system with $0 < a_1 < -1$ as shown in Fig. 3. However, because of the use of a "backwards" converter, this is similar to a "true" split. In this instance the split percentage is $a_1 = 0.57$. To make the system really workable, the converter has some unusual characteristics. The stator stage follows the impeller directly and this configuration makes for what is termed a "backwards" converter. This term is used to define a converter that has a complete reversal of rotation from impeller to turbine under normal operating conditions. The fluid vectors and configuration then result in an extremely high runout speed ratio (approximately 2.80 speed ratio), and a continuously rising torque absorption curve from stall to runout. When the package speed ratio varies from stall to 0.35 (in the hydraulic phase) the rotation of the sun gear (hence the impeller) is left-handed for a right-handed engine rotation. Thus, the torque flow diagram pictured in Fig. 3 is fulfilled since, in our convention, torque in the negative direction is no longer absorption but feedback to the planetary. The high speed ratios achieved by the turbine allow it to be ratioed down in speed and the torque multiplied before combining turbine torque to mechanical torque.

When the sun gear (impeller) is braked, the carrier rotates in a normal forward direction and an underdrive of 1.57/1. In this underdrive condition, the turbine is disconnected to minimize the hydraulic losses.

The actual performance curves for the Ilomatic package are shown in Fig. 35B. The engine lug in the hydraulic phase is not extreme, due to the shape

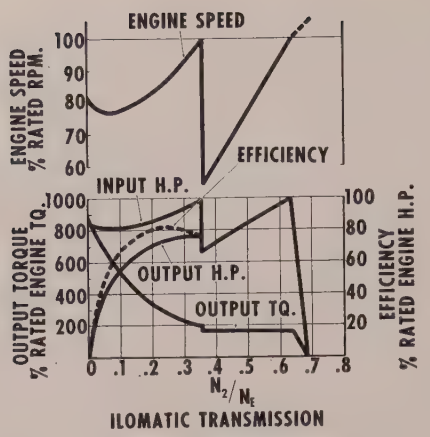


Fig. 35B — Performance characteristics, input split, Ilomatic transmission ($a_1 = -0.57$, $UR = 2.70$, $ER = 3.54$, $AR = 2.00$)

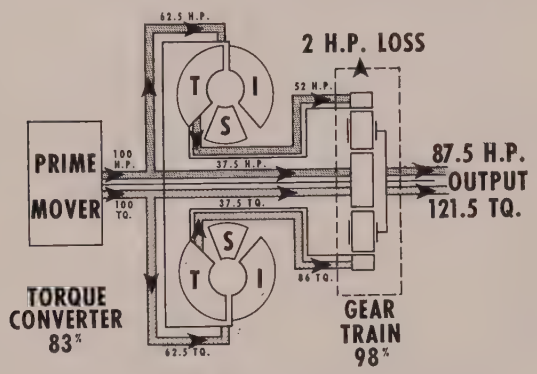


Fig. 36 — Output split — power and torque flow ($a_2 = 0.70$)

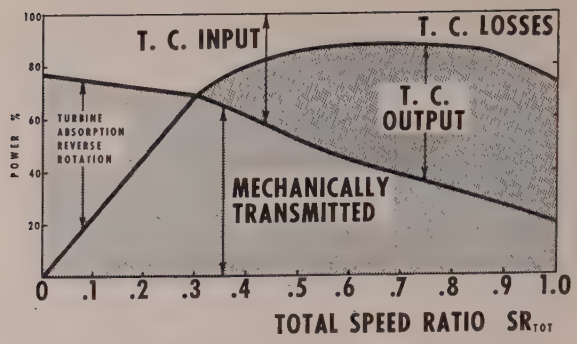


Fig. 37 — Output split — power analysis ($a_2 = 0.70$)

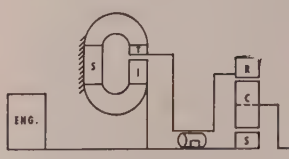


Fig. 38A — Output split schematic ($a_2 = 0.75$)

of the converter primary curve. The reason for an overall efficiency of approximately 82% with the split is that converter efficiencies of 80–90% are probably unattainable in this type of "backwards" converter. The output torque curve reaches some 850% of rated engine torque at stall but this occurs at the expense of a 0.64/1 torque ratio at full speed. The "actual" stall torque ratio at the output, when defined as the torque ratio increase from minimum output torque at full rated engine power in underdrive, is approximatly 5.3/1.

This unit serves to illustrate what wide possibilities remain in the recirculative systems. Some of these combinations and configurations might well make an excellent performing transmission, if the special converter characteristics necessary were produced with good efficiency.

True Output Splits

To cover completely the other "half" of these two-path split torque systems, we will now examine output split units with varying percentages of hydraulic split.

An output split where $a_2 = 0.70$ is pictured schematically in Fig. 36. The shaded lines indicate the power path through the unit at the speed ratio being examined. As in the previous illustration of an input split, the gearset is here credited with a 98% mechanical efficiency and the converter with 83% efficiency. The mechanical path pictured transmits 37.5% of the input power and the converter input absorbs 62.5% of the engine power. The overall efficiency of the package is then 87.5%. This shows a higher unit efficiency for the same percentage split than the input shown in Fig. 30. This is normal as an examination of the two power flows will quickly disclose. The general field of these efficiency comparisons was shown in Figs. 15 and 16.

The torque flow is shown in the dashed lines at the bottom half of the diagram. The torque values for the mechanical path and *hydraulic output* path show the 70% hydraulic-30% mechanical split.

A power analysis of this output split is shown in Fig. 37 for the complete range of unit speed ratios. The areas as labeled in this figure make the high speed ratio end of the curve quite clear. From 0.3 speed ratio back to stall we see occurring the negative speed ratio operation of the converter. This was shown in a general manner in Fig. 14. The power output to the converter impeller is determined strictly by the primary torque characteristics of the particular converter used. Some of the mechanically transmitted torque in this case feeds back into the turbine while turning it in the reverse direction. This actual curve of power absorption is somewhat different than the general curve shown in Fig. 13. The reason for the difference in these curves is because of the assumptions made of a straight-line torque absorption curve for the mythical converter used in Fig. 13.

The output split torque transmission pictured schematically in Fig. 38A is similar to that of Fig. 36. The only difference is the inclusion of the freewheel to lock the unit up at 1.0 speed ratio and to serve as a push-to-start device. The same reasons for the freewheel use are generally true in this unit as were true in the input split. However, with no freewheel in the output split unit, the no-load condition of the package will only drive the converter to its normal runout speed ratio. This would again cause a rise in the power dissipated to heat. The location of the freewheel is again a design variable as is its load. The freewheel in this type of instal-

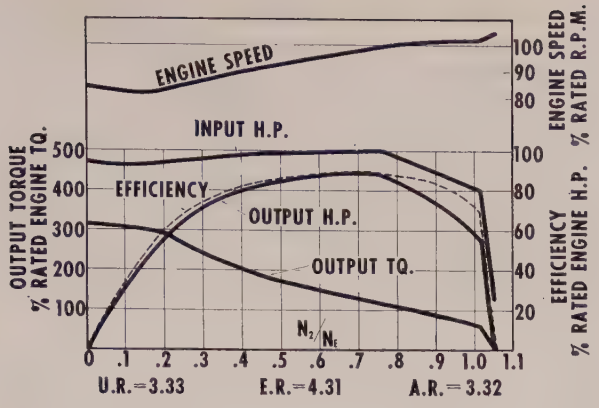


Fig. 38B — Performance characteristics output split ($a_2 = 0.75$)

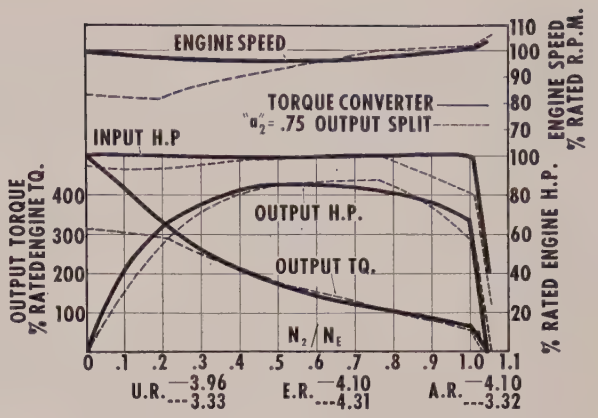


Fig. 39 — Performance comparison output split versus torque converter

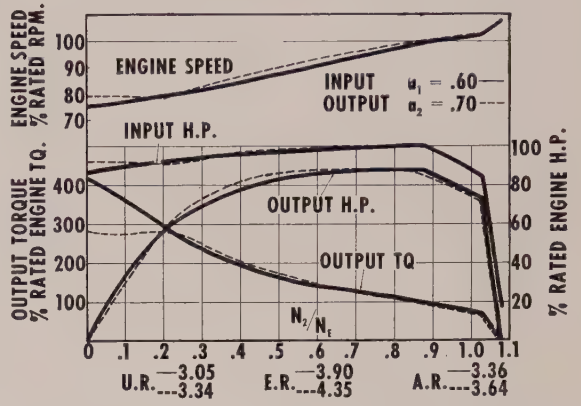


Fig. 40 — Performance comparison, input versus output split

lation needs a protective torque limiting device or it will be subjected to excessively high instantaneous torque peaks.

The performance curves for the unit shown in Fig. 38A are plotted in Fig. 38B. This type curve is a typical resultant output split curve when matched to a stationary housing single-stage converter. In this case, the engine is lugged to a maximum of 84% of full speed. The input horsepower curve shows that this particular engine is a rather high torque rise engine which helps keep the input power high even at 10–15% lug. The flattened output torque curve from 0.2 speed ratio down to stall is characteristic of output split units. It is caused by the converter operating in negative speed ratios. With converters that do not have steeply rising torque ratios with increasing negative speed ratios, the output torque curve may even drop from 0.2 down to stall. This would cause the actual stalled torque to be less than the output torque at some low output speed. This condition is not generally desirable but the flat curve as shown in Fig. 38B might be of some advantage to certain vehicle installations. This limiting of output torque as stall is approached might be advantageous for an installation that has severe track or wheel slippage at stall. A very high output torque at stall might be useless because of vehicle weight and traction. The input power at 1.05 output speed shows the heat required to be dissipated at no load.

The comparison of an output split torque package with a pure torque converter installation is made in Fig. 39. This comparison used a stationary housing single-stage converter of the same characteristics in both the split torque unit and the straight converter. The particular areas of difference are shown shaded and are especially evident in the engine speeds at stall, output torques at stall, peak output horsepower, and the heat absorption at 1.05 output speed. The comparative ratios are given at the bottom of the figure for both units. Note that the availability ratio is down while the economy ratio is up for the output split. The size of the converter used in the output split is 10% smaller than that used in the converter application. A comparison of two true splits input $a_1 = 0.60$ and output $a_2 = 0.70$ is made in Fig. 40. These units both used the same stationary housing single-stage converter and were matched to give nearly comparable output horsepower curves. The engine lug, input power, and output power are now very similar curves. There is a large difference in output torque at stall which is caused by the previously discussed negative speed ratio operation. The various ratios for these units are shown at the bottom of the figure. The output split unit in this case uses a 5% larger converter than the input split.

Another look at Figs. 8 and 16 shows the reasons that an output split $0 < a_2 < 0.5$ is not usually investigated for performance. Either one-half or more of the total speed ratio is with the converter in negative speed operation. This causes the efficiency curves to have a very high peak but be extremely narrow. Unless the application had another hydrodynamic unit as a starting device, the performance at stall would probably exclude use of these percentage splits.

Recirculative Output Splits

Figs. 41A and 41B are examples of a positive recirculative output split system with $a_2 = 3.25$. This system was portrayed schematically in Fig. 5. In this case, the hydrodynamic device is a fluid coupling. This unit is in production by the Twin Disc Clutch Co. under the name of "Hydro-Wynd." Its general application is for electric motor drives on machinery that requires constant tension on winding devices.

The high percentage hydraulic split (325%) requires the coupling to be somewhat larger than a pure coupling drive would necessitate. The split

arrangement allows the coupling to run at only 27% slip when the package is stalled. This greatly reduces the slope of the output torque curve and produces a rather flat output horsepower curve. The device, of course, has the other usual advantages of fluid coupling drive.

The split torque transmissions discussed herein have in general been assumed to have some sort of mechanical gear ratios behind them to form the total power transmission range for the vehicle. This would usually be in the form of a power shift transmission although not necessarily so. The analysis of the split torque package with added gear ratios is not made in this paper, since it would not contribute to the understanding of split torque transmissions.

The reason for using a stationary housing single-stage converter throughout most of the illustrations presented might now be discussed briefly. On all the "true" splits $0 < a < 1$ the package will usually increase impeller speed to some extent at stall (input splits) or move into negative speed ratio operation (output splits). This will cause engine lug, as has been illustrated. The use of a flat primary curve will minimize engine lug over the use of a converter that has a rising (1.0 to stall) primary curve. Therefore, the impeller primary torque characteristic of the stationary housing single-stage converter was an important reason for its use. Another factor was that the $0 < a < 1$ range will always cause some loss of package torque ratio. Therefore, a high torque ratio, high utility ratio converter is considered necessary to have some usable output torque left after combination with the gearing. This converter, with stall torque ratios in the range 4.5–6.5, fulfills this criterion.

The subject of the desirability of engine lug is rather interesting and ties in directly with split torque applications. Lug is desirable from a point of view of mechanical inertia if the transmission is so designed that the inertia can be transmitted to the wheels. Lug can be, and is, used for operator convenience in shifting by "ear" or by "smoke."

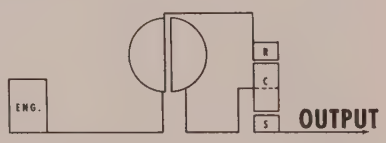


Fig. 41A — Output split with hydraulic coupling ($a_2 = 3.25$)

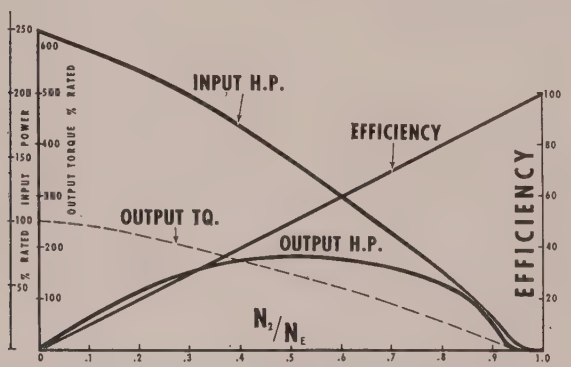


Fig. 41B — Performance characteristics, output split with coupling ($a_2 = 3.25$)

Noise control is also a factor at times. However, the psychology that demands that hydrodynamic-drive machines be made to lug, only because the usual operator "is used to" lug in his previous mechanical drive vehicle, is questioned. Engine lug is undesirable from a full throttle output horsepower point of view as a re-examination of Figs. 34 and 39 will show. An extreme lug will cause a drop in availability ratio. Generally, extreme engine lug is to be avoided for turbocharged engines or for operations that require a fast response time.

Results and Conclusions

For true split torque arrangements, the following advantages in comparison to straight converter drives may be listed:

1. The peak efficiency is increased.
2. The economy ratio is increased.
3. The converter size is reduced, down to 50%, depending on magnitude of split ratio and converter match point.
4. The engine speed is reduced as transmission stall is approached.
5. Some mechanical inertia effects are available which can be useful for various applications, but fluid drive smoothness is still maintained.
6. On sustained type converters, the power absorption at low-load high-speed conditions is reduced.
7. The planetary gear can be used for an overdrive or other purposes.

The disadvantages of the arrangement are:

1. The availability ratio is reduced.
2. The stall torque ratio is reduced.
3. There is less downhill braking available.
4. The design is more complicated, there are additional elements which compensate for the reduction of converter size.
5. The additional losses in gears and clutches may detract from the theoretical efficiency gains.
6. Torsional vibrations may occur in the drive line because there are additional elements connected to the engine flywheel which are not damped by the converter.

For recirculative systems, the following features are different:

Advantages — The stall torque ratio is increased, for output splits at least within certain limits.

Disadvantages —

1. The peak efficiency is reduced.
2. The economy ratio is reduced.
3. The converter size is increased.
4. The engine speed is reduced as high transmission speed ratios are approached.

Conclusions

Split torque arrangements are interesting for transmissions where:

1. Economy ratio is important and lower availability and stall torque ratios can be accepted (true splits).
2. Increase of stall torque ratio only is desired (positive recirculative systems).
3. Planetary of differential gears can be used for other purposes as well.

Split torque arrangements are not recommended for transmissions where:

1. Availability ratio is important.
2. Extreme simplicity of design is needed.

Discusser's Results Comparable With Authors'

— Ing. Hans Joachim Forster
Stuttgart-Unterturkheim — Bad Cannstatt

THE AUTHORS' RESULTS agree completely with those of my paper "Fottingergetriebe in Leistungsverzwiegungeh" (VDI-Forschungsheft 444).

In this paper I defined the converter speed ratio with a stalled third member of the planetary gearset as split ratio and designated it u_2 for the output split and u_1 for the inputs. In Bloch-Schneider the split ratio is a torque ratio. These ratios are both equivalent, and there are the relationships $a_1 = \frac{-u_1}{1-u_1}$ and $\frac{-1}{u_2-1}$. (There are minus signs because I use a minus sign for torques which have an opposite sense to the driving torques.) If one substitutes a_1 and a_2 by the corresponding expression for u_1 and u_2, the equations in both papers become identical.

But I recognize that the designation of a torque ratio as split ratio had this advantage; the same values of split ratios belong to true splits and recirculative systems for output as well as for input split torque arrangements.

Eq. 15 for the specific converter characteristic also corresponds to the assumption I used that the speed ratio of the design point becomes 0.5.

As far as the engine speed 4.1 is concerned, the only difference is that I plotted the engine speed versus output speed, whereas in the SAE paper, the engine speed is plotted versus total speed ratio (which in my figures was indicated as a parameter). Also, the equations for the converter size correspond if the expressions for the split ratios and for the converter characteristicc are modified. Since the two papers correspond to each other to such a degree it is not surprising that they lead to the same results and conclusions.

Because my paper deals mainly with the application of torque converters in the design of automobiles and trucks, I chose a stall-to efficiency range of 0.80 as the range of useful operation, and I compared the useful operation ranges of straight converters with those of split torque arrangements. Messrs. Bloch and Schneider publish other criteria. The assumption of the left border line at 70% efficiency or 70% power is preferable for industrial applications. The terms of economy ratio shows very well the influence of split torque arrangements upon efficiency. But one should not forget that there is a real improvement of transmission efficiency (efficiency plotted versus speed ratio) only to a minor extent. The major part of the increase in economy ratio arises because the converter operates for a longer period in the range of higher efficiency, generally the range of smaller torque ratio.

It should also be mentioned that the arrangements with negative split ratios as used by the Ilomatic transmission belong to the most interesting configurations. I have pointed this out in my paper in Sec. 6.43 and Figs. 42 and 44. This kind of arrangement increases the efficiency and the stall torque ratio, but reduces substantially the transmission speed ratio.

The torque converter used in the Ilomatic transmission has, therefore, a very high speed ratio (as indicated by the authors) with an accordingly low efficiency. If this kind of split torque arrangement is combined with this kind of converter, we obtain again a transmission whose characteristics are not too different from those of a normal, pure converter.

The very impressive comparisons between straight converter drives and split torque arrangements in the Figs. 32B, 34, 39B, 39 and 40 show that a split torque design really leads to only minor modifications of the transmission characteristics and that it reduces the useful ranges when it increases the efficiency.

With the exception of some special cases, improvement of efficiency as well as extension of the useful range constitute the objectives of development. For that reason the increased complication of split torque arrangements will only be justified under certain conditions. In transmission arrangements with multiple ranges or speeds which must be shifted, the use of split torque arrangements can sometimes improve the shifting characteristics — because sharp reductions of the output torque may be avoided.

80% Split Torque Units Have Numerous Advantages

— C. A. Ramsel
Caterpillar Tractor Co.

IF WE LIMIT our discussion to true splits, both input and output, and compare characteristics of the units at the 80% points instead of the 70% points, we arrive at a few different conclusions than those presented by the authors.

The economy ratios using 80% points are very much improved in the split units. This higher efficiency throughout the entire range materially reduces the heat rejected to coolers. The utility and availability ratios are definitely improved at these 80% points. Further improvement can be had by using engines with high torque rise characteristics. We feel that these 80% points are realistic since they cover a broad enough speed range to allow transmissions with step ratios of approximately 2/1 if necessary.

The reduction in stall torque ratio is appreciated in some respects. A machine is geared for a particular performance at a finite speed rather than at stall conditions, therefore, a flatter torque curve permits the use of lower capacity clutches to hold converter stall and reduces shock while shifting.

An extended low-load high-speed range of an extended low-speed high-torque range at low efficiency are not considered desirable for vehicular operation. A split torque drive generally has less torque rise between 80% efficiency and stall and less speed increase between 80% efficiency and no load than does a straight converter.

These remarks are in reference to power shifted machines and the qualities we find desirable for them.

Suggests Gear Hookup To Eliminate One Ratio

— Adiel Y. Dodge
A. Y. Dodge Co.

WHY NOT USE a gear hookup which provides a regenerative or feedback at the start, one which changes to a 100% hydraulic drive during intermediate operation and to a nonregenerative two-path torque split at low torque?

As brought out in Figs. 5 (upper) and 15, the regenerative split moves the efficiency curve to the left and as brought out in your Figs. 36 and 16, an output split lifts efficiency curve and moves it to the right. It is submitted that such a combination as herein suggested will not only improve the utility ratio and the extent of the economy ratio (particularly above 75% efficiency), but it will also improve the performance at the start.

My conclusion is based on the authors' examples and figures, not necessarily the most favorable selection. Bear in mind that all vehicles must start at zero efficiency, and that there are work vehicles which make many starts and stops.

In my opinion, such an arrangement used in work vehicles will make it possible to eliminate one mechanical ratio in the step transmission employed behind the torque converter planetary gear arrangement. The elimination of the one step will approximately compensate for the added parts required to follow my suggestion.

This arrangement has other advantages. It renders an advantage in the selection of the torque converter type and size as well as increased performance to move out of the lower into the higher efficiency range in less time. This also improves the work factor, also reduces engine lugging.

Further information on this suggestion will be furnished upon request.

application

Cumulative Fatigue

to practical

W. F. Hofmeister
Chain Belt Co.

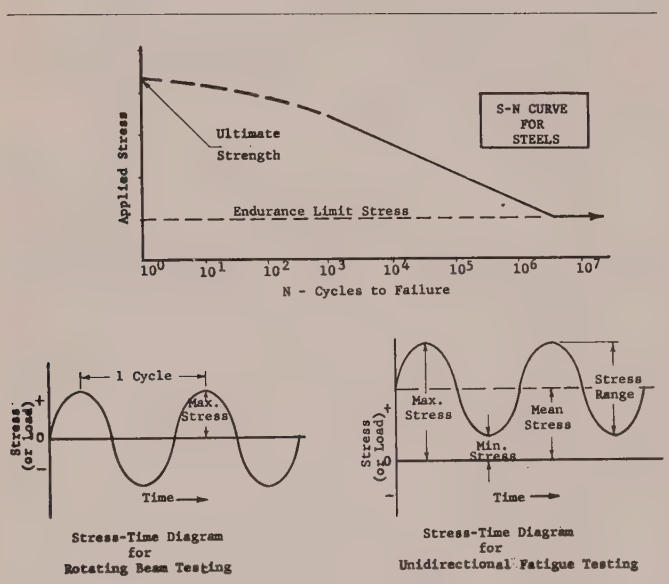


Fig. 1 — Steady-state cyclic stressing

SINCE THE TIME it was first recognized that "brittle" failures in parts subjected to cyclic stresses were not the result of "crystallization" of the metal, but were a basic strength limitation, much has been done to determine and catalog the fatigue strength of engineering materials. In more recent years, manufacturers have been conducting exhaustive tests on their products to establish experimentally the true fatigue strength of the parts as manufactured. The results of these tests made it apparent that while the data were useful in improving product quality, they were of only limited value to design engineers. In the few cases where parts were designed to be applied below their endurance limit at the maximum service loads imposed, the data could be applied directly to design problems. However, the test results also showed that they had been using materials successfully at well above their endurance limit. Designers required information that could be used to predict the life of the part when it was loaded so that endurance limit stresses were exceeded, and were faced with the further complication that service loads were completely heterogene-

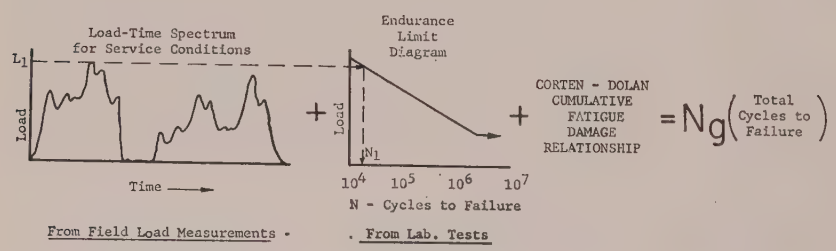


Fig. 2 — Elements required to predict life of parts subjected to heterogeneous load spectrum

of the

Damage Theory

problems

THIS PAPER presents an analysis of the Corten-Dolan cumulative fatigue damage theory. This equation was used to predict total cycles to failure for random dynamic loading on chains operating over sprockets in the laboratory.

This theory takes into account all peak stresses to which the part is subjected. And it assumes that the various stress ranges are not applied in any given sequence.

The author also describes various other methods for predicting cycles to failure for steels subjected to varying load ranges.*

ous rather than the uniform cyclic loads imposed in fatigue tests. The discussion following covers a method of bridging the gap between the data obtained in tests conducted at constant stress amplitude, and the fatigue life of a part subjected to the heterogeneous loads experienced in service.

Throughout this discussion, stresses below those required to cause gross plastic deformation will be considered. Under these conditions, stress in the region of failure is proportional to the applied load. Hence, the terms stress and load will be used interchangeably as the occasion demands.

Fatigue Problem — Repetitive Stressing

Fig. 1 shows the relationship between the applied stress and the number of times this stress can be applied and removed before the part fails. Failure is defined as complete rupture of the part. Before rupture occurs, a fatigue crack is always present. Before the fatigue crack can be observed, using even microscopic methods of examination, damage to the material has occurred. The number of cycles between the first appearance of the crack and final rupture is generally small compared to the total cycles to failure. Hence, using complete rupture as the failure criterion is sufficiently precise for design purposes.

Fig. 1 shows also two common means of loading employed in fatigue tests. The lower left diagram shows the stress-time history for reversed bending (Moore Rotating Beam) tests. The lower right diagram shows the type loading generally applied in tests on actual parts or structures. Also, except for the fact that the amplitude of the stresses varies from cycle to cycle, this is the type of loading most often encountered in service. The data on chain used in the problem to be described later were obtained in fatigue tests employing this type of loading.

Fig. 2 illustrates the system discussed in this paper for predicting the fatigue life of parts subjected to varying loads greater than their endurance limit. The two elements on the left in Fig. 2 must be obtained by direct testing. Several theories have been proposed relating to the element concerning the "Cumulative Fatigue Damage Relationship."[1,2,3,4] The one being considered in this paper is the relationship developed by Corten and Dolan.[4] It differs from the expressions developed by others in that it considers the propagation, at stresses below the endurance limit, of fatigue damage induced by prior operation at high stress levels. As will be shown later, the predictions made using the Corten-Dolan relationship are in good agreement with the service history of a chain application.

Effect of Stress Range on Cycles to Failure

Fig. 3 shows the Modified Goodman Diagram method for predicting the endurance limit of a material subjected to any load range when the endurance limit at zero minimum load, or zero mean load, is known. Data developed in the Chain Belt Co. Laboratory to be discussed later show that the dashed limited life lines in Fig. 3 give conservative

* Paper presented at meeting of SAE Milwaukee Section, May 1, 1959.

[1] "Effect of Overstress in Fatigue on Endurance Life of Steel," by J. B. Kommers. ASTM Proceedings, Vol. 45, 1945.

[2] "Fatigue and Hardening of Steels," by H. J. French. Transactions, American Society Steel Treating, Vol. 21, 1933.

[3] "Cumulative Damage in Fatigue," by M. A. Miner. ASME Transactions, Vol. 67, 1945.

[4] "Cumulative Fatigue Damage," by H. T. Corten and T. J. Dolan. Paper presented at International Conference on Fatigue of Metals, New York, November, 1956 (paper 2, session 3).

predictions for intermediate hardness quenched and tempered steels. Reviews of the data available in the literature on fatigue tests conducted at various load ranges were published by J. O. Smith.[5] Fig. 4A shows these data as plotted by Smith. It is apparent that for high hardness (relatively brittle) steels, the Modified Goodman Diagram method predicts an endurance limit at any load range which is in reasonable agreement with the experimental results. For ductile steels, however, this method leads to highly conservative results. For ductile steels, mean load has no significant effect on the endurance limit load range for mean loads up to about 30% of the ultimate strength of the material. This suggests revision of the Goodman Diagram to that shown in the lower sketch in Fig. 4B. Up to 30% of the ultimate strength of the material, the minimum and maximum load lines are parallel for both the endurance limit and the finite life load ranges.

Fatigue tests on chain at various load ranges were conducted in the chain Belt Co. Laboratory on the Schenck Fatigue Machine shown in Fig. 5. Seven pitch strands are generally used, hence experimental scatter is greatly reduced, compared to tests on single samples, since the failures occur at the pin or bushing holes and each failure represents a "minimum of 16" or "minimum or 12" potential points of failure. The data obtained for minimum loads ranging from about 3 to 50% of the ultimate strength are given in Fig. 6. It is evident that when the load plotted is that obtained using the Modified Goodman method, actual cycles to failure is greater than for samples on which the minimum load was maintained at nearly zero. If cycles to failure is plotted against the "load range" for the particular test, actual cycles to failure is less on samples run at high minimum loads than for samples on which the minimum load was maintained at nearly zero.

The further comparisons of these data given in Fig. 7 show that for the steels used in chain sideplates (about 400 Bhn), the arithmetic average of the load calculated using the Goodman hypothesis and the load range predicts cycles to failure with an accuracy which is within the scatter band of the data itself.

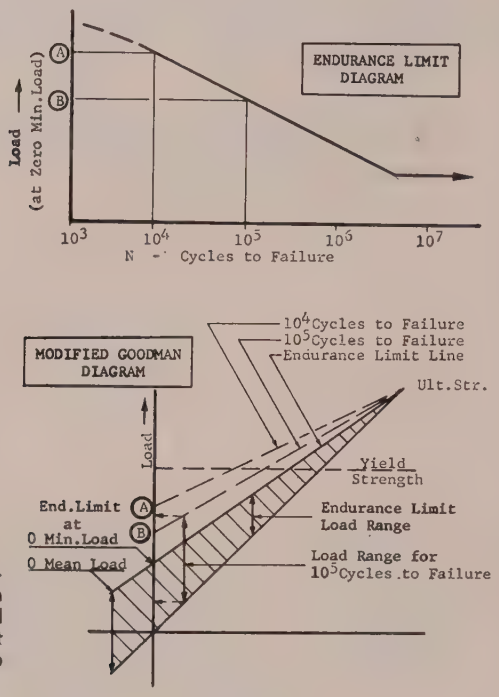
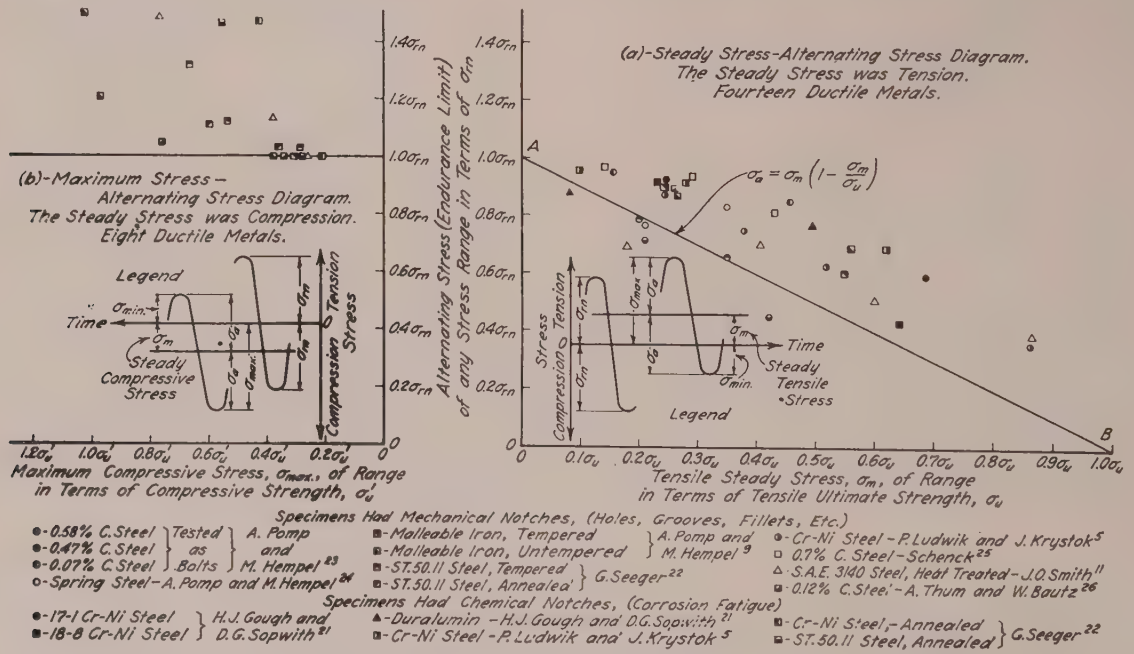


Fig. 3 — Modified Goodman Diagram method for limited life at other than zero mean stress

[5] "Effect of Range of Stress on Fatigue Strength of Metals," by J. O. Smith. Bulletin 334, Engineering Experiment Station, University of Illinois, 1942.
[6] "Investigation of Fatigue of Metals," by H. F. Moore and J. B. Kommers. Bulletin No. 124, Engineering Experiment Station, University of Illinois, 1921.


(a)-Steady Stress-Alternating Stress Diagram. The Steady Stress was Tension. Fourteen Ductile Metals.
(b)-Maximum Stress – Alternating Stress Diagram. The Steady Stress was Compression. Eight Ductile Metals.
Legend
Time
Steady Compressive Stress
Steady Tensile Stress
Compression
Tension
Alternating Stress (Endurance Limit) of any Stress Range in Terms of σ_{rn}
Maximum Compressive Stress, $\sigma_{max.}$, of Range in Terms of Compressive Strength, σ'_u
Tensile Steady Stress, σ_m, of Range in Terms of Tensile Ultimate Strength, σ_u
$\sigma_a = \sigma_m \left(1 - \frac{\sigma_m}{\sigma_u}\right)$
Specimens Had Mechanical Notches, (Holes, Grooves, Fillets, Etc.)
0.58% C. Steel
0.47% C. Steel
0.07% C. Steel
Spring Steel – A. Pomp and M. Hempel
Tested as Bolts
A. Pomp and M. Hempel
Malleable Iron, Tempered
Malleable Iron, Untempered
A. Pomp and M. Hempel
ST. 50.11 Steel, Tempered
ST. 50.11 Steel, Annealed
G. Seeger
Cr-Ni Steel – P. Ludwik and J. Krystok
0.7% C. Steel – Schenck
S.A.E. 3140 Steel, Heat Treated – J.O. Smith
0.12% C. Steel – A. Thum and W. Bautz
Specimens Had Chemical Notches, (Corrosion Fatigue)
17-1 Cr-Ni Steel
18-8 Cr-Ni Steel
H.J. Gough and D.G. Sopwith
Duralumin – H.J. Gough and D.G. Sopwith
Cr-Ni Steel – P. Ludwik and J. Krystok
Cr-Ni Steel, Annealed
ST. 50.11 Steel, Annealed
G. Seeger

Range of Stress Diagrams for Notched Specimens of Ductile Metals Subjected to Ranges of Repeated Axial Stress Superimposed on a Steady Stress

Fig. 4A — Range of stress diagrams for notched specimens of ductile metals subjected to ranges of repeated axial stress superimposed on steady stress

Considering the foregoing comparisons, the methods for predicting cycles to failure for steels subjected to varying load ranges shown in Table 1 appear to give the best correlation with the experimental evidence.

Cycles to Failure for More than One Stress Range

Early approximations of the influence of subjecting a part to several stress ranges on total cycles to failure are shown in Fig. 8. These predictions were based on the hypothesis that each cycle at a given over stress caused $1/N_i$ of the total damage at N_i cycles, where N_i is the cycles to failure at the given stress. Kommers and Moore[6] proposed this hypothesis for a part subjected to two discrete stress ranges. Minor[3] expanded it further to encompass any number of stress ranges. Considerable work by

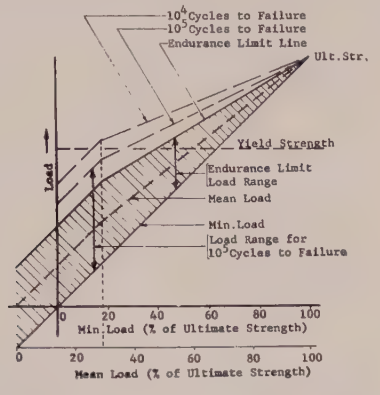


Fig. 4B — Modified Goodman Diagram for ductile steels (from summary by J. O. Smith)

Table 1 — Methods for Predicting Cycles to Failure for Steels

Hardness Range (Bhn)	Method Giving Most Accurate Prediction of Fatigue Life as Determined at Zero Minimum Load
Less than 203	Load Range
203–463	$\frac{\text{Load Range} + \text{Modified Goodman Diagram}}{2}$
Greater than 463	Modified Goodman Diagram

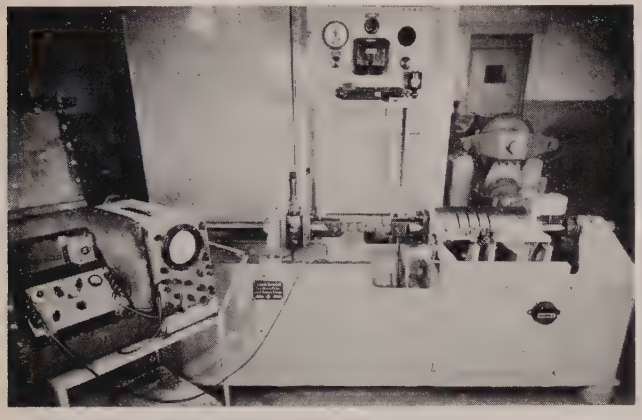


Fig. 5 — Six-ton (metric) Schenck fatigue machine

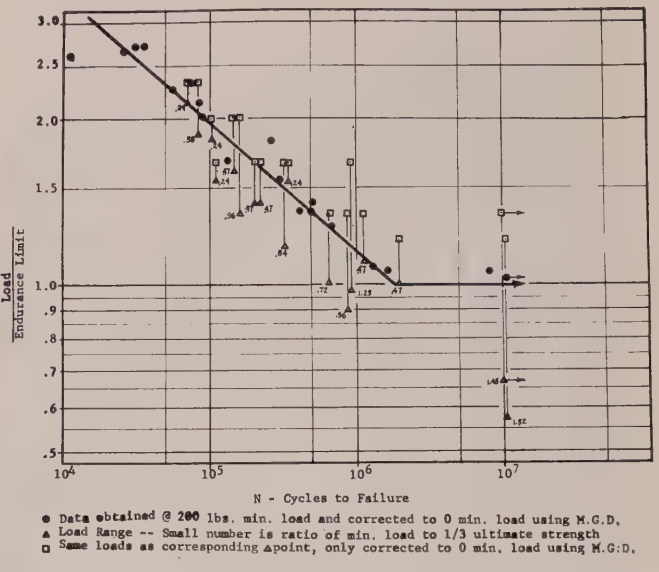

Load / Endurance Limit
N - Cycles to Failure
● Data obtained @ 200 lbs. min. load and corrected to 0 min. load using M.G.D.
▲ Load Range -- Small number is ratio of min. load to 1/3 ultimate strength
□ Same loads as corresponding ▲point, only corrected to 0 min. load using M.G.D.

Fig. 6 — Effect of load range on fatigue life of ASA roller chain

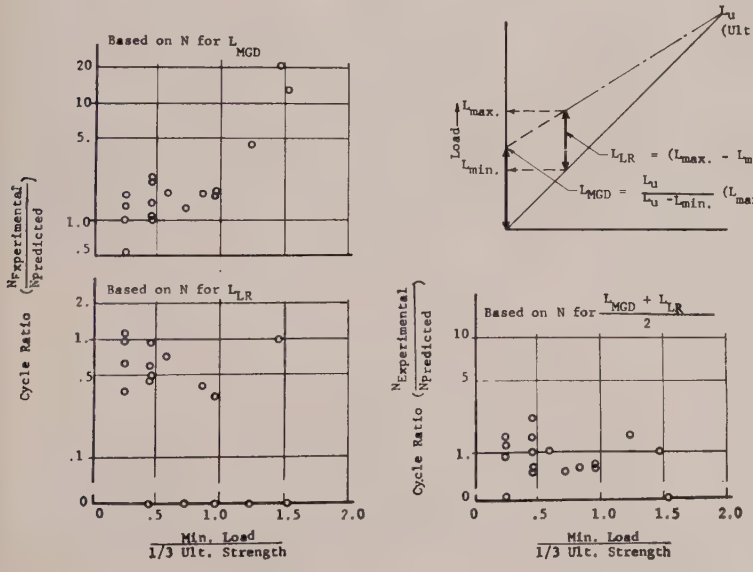


Fig. 7 — Comparison of methods for predicting fatigue life for various load ranges on ASA roller chains

many investigators[1,7,8,9,10,11] showed that it made appreciable difference to total fatigue life whether the initial cycles were applied at a greater or lesser stress level than the final cycles to failure. The results of these investigations, given qualitatively in Fig. 9, showed that when the initial stress was greater than the final test stress, per cent damage per cycle was greater than when the initial stress was lower than the final stress. While other methods for predicting fatigue life for random loading have also been proposed,[12] the remainder of this paper will be devoted to the equation developed by Corten and Dolan.[4]

The equation is:

$$N_g = \frac{N_1}{\alpha_1 + \alpha_2 \left(\frac{\sigma_2}{\sigma_1}\right)^a + \alpha_3 \left(\frac{\sigma_3}{\sigma_1}\right)^a + \ldots + \alpha_i \left(\frac{\sigma_i}{\sigma_1}\right)^a}$$

where:

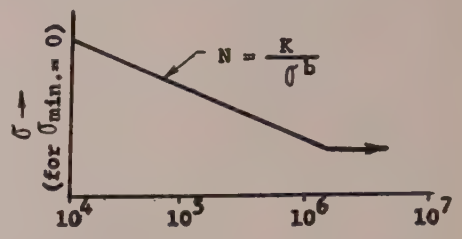
(Answer desired) N_g = Total cycles to failure
(From S-N diagram) N_1 = Cycles to failure for continuous stressing at σ_1

Data from knowledge of stressing (or loading) encountered in service:

σ_1 = Maximum applied stress
σ_2 = Second largest applied stress
σ_3 = Third largest applied stress
σ_i = Minimum applied stress at which damage progresses
α_1 = Ratio of cycles n_1 at σ_1 to total cycles N_g
α_2 = Ratio of cycles n_2 at σ_2 to total cycles N_g
α_i = Ratio of cycles n_i at σ_i to total cycles N_g
$a = C \times b$ Note: The experimental work done at the University of Illinois showed that a value of C between 0.81 and 0.94 resulted in close predictions of total cycles to failure for steel specimens subjected to varying load ranges.

From S-N diagram:

$$b = \frac{K' - \log N}{\log \sigma} \qquad (K' = \log K)$$



Note: The diagram above shows that to determine b, only the limited life portion of the endurance limit diagram needed be established. The latter requires relatively little testing time compared to establish-

---

[7] "Der Einflus Einer Wechselvorbeanspruchung auf Biegezeit und Biegewechselfestigkeit von Stahl St. 37," by H. Müller-Stock, E. Gerold, and E. H. Schulz. Archiv für Eisenhüttenwessen, Vol. 12, 1938.
[8] "Hypothesis for Determination of Cumulative Damage in Fatigue," by F. E. Richart and N. M. Newmark. ASTM Proceedings, Vol. 48, 1948.
[9] "Concept of Fatigue Damage," by S. M. Marco and W. L. Starkey. ASME Transactions, Vol. 76, 1954.
[10] "Study of Damaging Effect of Fatigue Stressing of X4130 Steel," by J. A. Bennet. ASTM Proceedings, Vol. 46, 1946.
[11] "Effect of Prior Repeated Stressing on Fatigue Life of 75S-T6 Aluminum," by T. J. Dolan and H. F. Brown. ASTM Proceedings, Vol. 52, 1952.
[12] "Design for Fatigue," by J. Marin. *Machine Design*, Feb. 21, 1957.

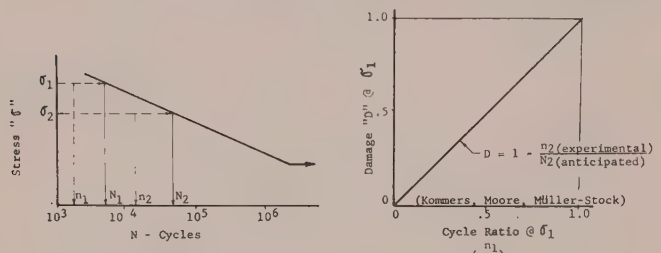

Stress "σ"
N - Cycles
Damage "D" @ σ1
Cycle Ratio @ σ1
(Kommers, Moore, Müller-Stock)
D = 1 - n2(experimental)/N2(anticipated)
Σ n1/N1 + n2/N2 + ... + ni/Ni = 1 (Miner)
Where: ni = Cycles applied @ σi
Ni = Cycles to failure @ σi
Nanticipated = N2 Cycles to failure for σ2 on a virgin specimen.

Fig. 8 — Prediction of fatigue life for parts subjected to more than one stress range

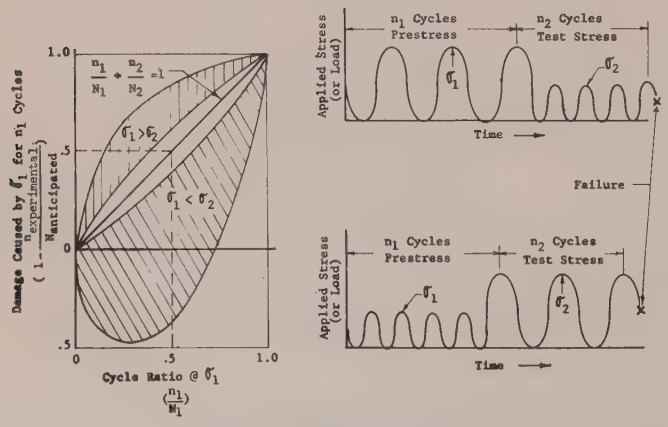


Fig. 9 — Effect of relative magnitude of pre-stress and test stress on total fatigue life (from Müller-Stock, Kommers, Bennet, Richart & Newmark, Marco & Starkey, Brown & Dolan)

$$1 = \frac{n_1}{N_1} + \frac{n_2}{N_2} + \frac{n_3}{N_3} + \ldots + \frac{n_i}{N_i} \quad \text{(Miner)}$$

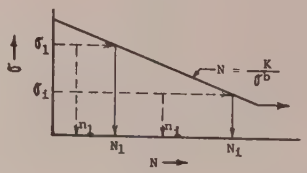
Let:

N_g = Total cycles to failure
N_1 = Cycles to failure for σ_1 applied continuously

$$\frac{N_1}{N_g} = \frac{n_1}{N_g} \times \frac{N_1}{N_1} + \frac{n_2}{N_g} \times \frac{N_1}{N_2} + \frac{n_3}{N_g} \times \frac{N_1}{N_3} + \ldots + \frac{n_i}{N_g} \times \frac{N_1}{N_i}$$

Let: $\alpha_1 = \frac{n_i}{N_g}$ = Frequency ratio

$$N_g = \frac{N_1}{\alpha_1 \left(\frac{N_1}{N_1}\right) + \alpha_2 \left(\frac{N_1}{N_2}\right) + \alpha_3 \left(\frac{N_1}{N_3}\right) + \ldots + \alpha_1 \left(\frac{N_1}{N_i}\right)}$$



$$\frac{N_1}{N_2} = \frac{\frac{K}{\sigma_1^b}}{\frac{K}{\sigma_2^b}} = \left(\frac{\sigma_2}{\sigma_1}\right)^b$$

$$N_g = \frac{N_1}{\alpha_1 + \alpha_2 \left(\frac{\sigma_2}{\sigma_1}\right)^b + \alpha_3 \left(\frac{\sigma_3}{\sigma_1}\right)^b + \ldots \alpha_i \left(\frac{\sigma_i}{\sigma_1}\right)^b}$$

For Miner Hypothesis
σ_i (min) = Endurance limit
b = Slope of endurance limit diagram
For Corten-Dolan
Accounts for propagation of, at low stress, damage induced by prior operation at high stress.
σ_i (min) = Lower than endurance limit
$b = 0.85 \times$ Slope of endurance limit diagram

Fig. 10 — Development of Corten-Dolan relationship and comparison with Miner hypothesis

ing the endurance limit. Testing must be conducted to at least 10 million cycles to have reasonable confidence that an endurance limit has been obtained.

It can be seen that this expression takes into account all peak streses to which the part is subjected, that is, σ_1, σ_2, σ_3, and the rest. It is assumed that the various stress ranges are *not* applied in any given sequence, that is, the order of application is random. Random loading has been found to be the case for aircraft structures. It was also found to be the case in the chain application to be considered in the remainder of this paper. It is assumed to be true for most operating machinery and vehicles. It will be shown that the neglecting stress ranges where σ_i/σ_1 is less than 0.1 has no significant effect on the accuracy of the prediction.

H. T. Corten has shown that the size of the stress range increment into which the data are grouped has only a minor influence on the cycles to failure calculated using the Corten-Dolan equation. As the size of the increment increases, the number of terms in the denominator on the right side of the equation decreases, but the magnitude of each term increases. Thus, the summation of terms changes very little, and the difference in prediction resulting from dividing the total dynamic stress spectrum into 5 or 10 increments is smaller than the confidence level that can be assigned to the repeatability of the dynamic stress spectrum in practical problems.

If the exponent a on the intensity factors in the Corten-Dolan equation were equal to the slope of the limited life line b from the S-N diagram, the Corten-Dolan and Miner equations can be shown to be identical (Fig. 10). Thus, the only differences between the two equations are:

1. Each term in the denominator is larger in the Corten-Dolan equation ($a \cong 0.85\ b$).

2. More terms are included in the denominator of the Corten-Dolan equation (Miner's relationship neglects the damage propagated by stresses below the endurance limit of a virgin specimen, that is, as shown in Fig. 8, N_i will be infinite and n_i/N_i will be

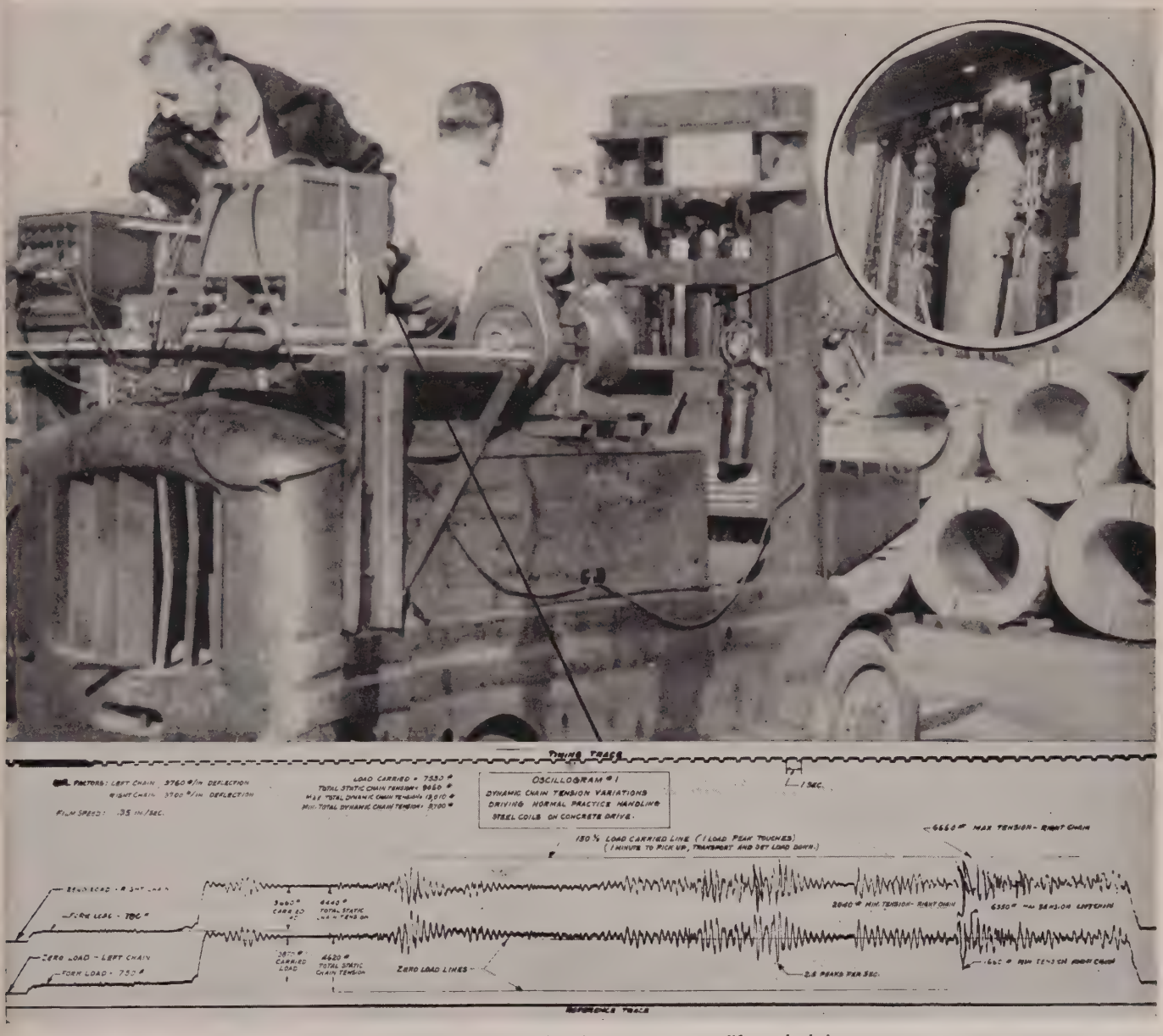

OSCILLOGRAM #1
DYNAMIC CHAIN TENSION VARIATIONS
DRIVING NORMAL PRACTICE HANDLING
STEEL COILS ON CONCRETE DRIVE.
1 SEC.
150% LOAD CARRIED LINE (1 LOAD PEAK TOUCHES)
(1 MINUTE TO PICK UP, TRANSPORT AND SET LOAD DOWN.)
6660# MAX TENSION - RIGHT CHAIN
2040# MIN. TENSION - RIGHT CHAIN
ZERO LOAD - LEFT CHAIN
FORK LOAD - 750#
ZERO LOAD LINES
2.6 PEAKS PER SEC.

Fig. 11 — Test setup for field load measurements on lift truck chains

Table 2 — Summary of Data from Field Load Measurements on Lift Truck

Load Carried by Truck, lb	5001–6000	6001–7000	7001–8000	8001–9000	9000 +	
Approximate No. of Carries per Day	10	20	20	20	10	80 Total/8-hr day
Maximum Dynamic Chain Tension, lb	6000	6700	7200	7800	8200	

Maximum Dynamic Tension, lb	Dynamic Tension ranges, lb	Relative Frequency of Given Chain Tension Range										Σ of Normalized Cycles	Frequency Factor $a_i = \frac{\text{Cycles at } T_i}{\Sigma \text{ Cycles}}$
		From Graph	Normal. to 1240	From Graph	Normal. to 2480	From Graph	Normal. to 2480	From Graph	Normal. to 2480	From Graph	Normal. to 1240		
8200		—	—	—	—	—	—	—	—	10	10	10	$a_1 = 0.0010$
7800		—	—	—	—	—	—	20	20	—	—	20	$a_2 = 0.0020$
7200		—	—	—	—	20	20	—	—	—	—	20	$a_3 = 0.0020$
6700		—	—	20	20	—	—	—	—	—	—	20	$a_4 = 0.0020$
6000		10	10	—	—	—	—	—	—	—	—	10	$a_5 = 0.0022$
	6000	0	0	0	0	0	0	0	0	1	12	12	
	5000	2.8	34	6.4	69	6.4	69	6.4	69	2.8	34	275	$a_6 = 0.0275$
	4000	8	98	17	183	17	183	17	183	8	98	745	$a_7 = 0.0745$
	3000	22	271	50	538	50	538	50	538	22	268	2153	$a_8 = 0.2153$
	2000	60	739	140	1507	140	1507	140	1507	60	730	5990	$a_9 = 0.5990$
	1000	8	98	17	183	17	183	17	183	8	98	745	$a_{10} = 0.0745$
	Totals	100.8	1250	230.4	2500	230.4	2500	230.4	2500	101.8	1250	10,000	1.0000

Table 3 — Chain Life Calculation Using Corten-Dolan Relationship

$$N_g = \frac{N_1}{a_1 + a_2\left(\frac{T_2}{T_1}\right)^a + a_3\left(\frac{T_3}{T_1}\right)^a + \ldots + a_i\left(\frac{T_i}{T_1}\right)^a}$$

Dynamic Tension Range, lb	a_i (from Table 2)	$\frac{T_i}{T_1}$	$\left(\frac{T_i}{T_1}\right)^{3.47}$	$a_i \times \left(\frac{T_i}{T_1}\right)^{3.47}$
8200	0.0010	1.000	1.000	0.00100
7800	0.0020	0.951	0.840	0.00168
7200	0.0020	0.879	0.639	0.00128
6700	0.0020	0.817	0.496	0.00099
6000	0.0022	0.732	0.339	0.00075
5000	0.0275	0.610	0.180	0.00495
4000	0.0745	0.488	0.083	0.00619
3000	0.2153	0.366	0.0306	0.00659
2000	0.5990	0.244	0.0075	0.00449
1000	0.0745	0.122	0.0007	0.00005
				0.02797

N_1 for 8200 lb Maximum Load $= 23 \times 10^3$ Cycles

$$N_g = \frac{N_1}{\Sigma a_i \left(\frac{T_i}{T_1}\right)^a} = \frac{23 \times 10^3}{2.797 \times 10^{-2}} = 820{,}000 \text{ Cycles to failure}$$

zero for all σ_i, less than the endurance limit).

Both differences result in Miner's equation predicting greater cycles to failure than are predicted by the Corten-Dolan relationship.

Application of Cumulative Fatigue Damage Theory to Chain Problem

Fig. 11 shows a lift truck and the instrumentation used to measure the dynamic tensions imposed on the chains which raise the forks. A typical record showing the cyclic chain loading encountered in raising, transporting, and lowering the load carried is included in Fig.11. In this particular application, the truck was used to lift known heavy loads continuously and to transport them over the same driveway. Chain failures had been occurring every 3–4 months. This made it an ideal application on which to test the limited life predictions of the Corten-Dolan cumulative fatigue damage relationship. The problem was first studied by H. T. Corten. The analysis following is essentially the same as presented in T. & A.M. Report No. 538, University of Illinois. It differs in the interpretation given to the maximum dynamic loads measured, and in that additional information was obtained regarding the work done by the truck.

The number of cycles of each dynamic chain tension range counted on the records obtained for typical operating conditions was recorded, and the data grouped into 500-lb tension range increments. Because the minimum loads were very low compared to the ultimate strength of the chain, tension range as read off the oscillograms could be used directly without introducing significant errors. The number of cycles in each tension range increment for the various loads carried by the truck was normalized to result in the following conditions:

1. One cycle of the maximum tension range recorded.

2. Twice as many cycles for the intermediate loads carried compared to the maximum and minimum loads carried, the condition production records showed prevailed in service.

To average the data and minimize experimental scatter, the normalized data were plotted as shown in Fig. 12. It is evident in this figure that while the greatest dynamic tension range occurred when the lift truck transported the largest load, the number of cycles for each tension range below about 5000 lb is nearly independent of the load carried by the

Table 4 — Comparison between Miner and Corten-Dolan Life Predictions

Dynamic Tension Range, lb	Corten-Dolan (from Table 3) $a_i \times \left(\frac{T_i}{T_1}\right)^{3.47}$	Miner $a_i \times \left(\frac{T_i}{T_1}\right)^{4.09}$
8200	0.00100	0.00100
7800	0.00168	0.00163
7200	0.00128	0.00118
6700	0.00099	0.00087
6000	0.00075	0.00062
5000	0.00495	0.00363
4000	0.00619	———
3000	0.00659	———
2000	0.00449	———
1000	0.00005	———
	0.02797	0.00893

$\frac{0.02797}{0.00893} = 3.14$ Longer life predicted by Miner Hypothesis

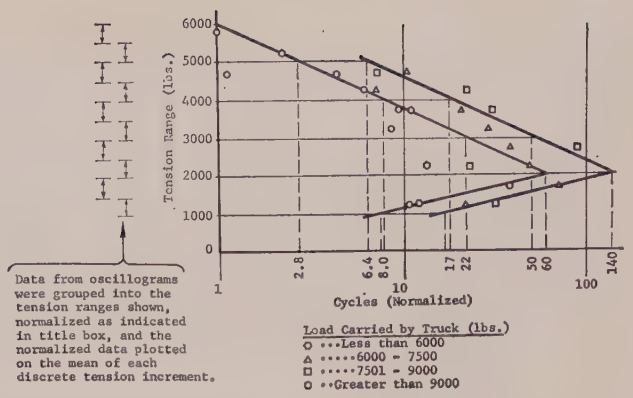


Fig. 12 — Plot of number of cycles of each tension range to which the chains were subjected (data from oscillograms of dynamic chain loading). Number of cycles counted were normalized to result in: (1) one cycle of maximum tension range, (2) twice as many cycles for intermediate loads carried by truck as for maximum and minimum loads carried

truck. The two curves are displaced because the heaviest and lightest loads are carried half as often. While no data were obtained for transporting loads of 7001–8000 lb, it can be assumed that the tension ranges imposed on the chains would be the same as when the 6001–7000 lb and the 8001–9000 lb loads were transported.

Examination of the oscillograph records showed that approximately 125 cycles of loading occurred on the chains each time the truck was used to lift, transport, and deposit the load. For the estimated total of 80 lifts per day, this results in 10,000 cycles of dynamic loading per day. The data from Fig. 12 were re-normalized as shown in Table 2 to give the estimated number of cycles of each tension range occurring per day. The frequency ratios α_i, and the corresponding intensity ratios T_i/T_1, are given in Table 3.

$$\text{Days life} = \frac{8.2 \times 10^5 \text{ cycles to Failure}}{10^4 \text{ cycles per Day}} = 82 \text{ Days to failure}$$

$$\text{Months life} = \frac{82 \text{ days to Failure}}{21 \text{ working Days per Month}} = 3.9 \text{ Months to failure}$$

The exponent a applied to the T_i/T_1 ratios was obtained by multiplying by 0.85 the b from the equation of the Endurance Limit Diagram of the chain in question, that is:

$$N = \frac{K}{T^b} \text{ (Determined by Chain Belt Co.)}$$

$$T = \text{Tension range (lb)}$$

$$a = 0.85 \times b = 0.85 \times 4.09 = 3.47$$

The number of cycles to failure at the maximum chain tension measured was also obtained from the Endurance Limit Diagram for this chain: $N_1 = 23{,}000$ cycles to failure at 8200 lb maximum and 0 lb minimum tension.

The $\alpha_i \times (T_i/T_1)^a$ products given in the last column of Table 3 are proportional to the fatigue damage produced for the cycles of operation at each of the conditions tabulated. Since these products are greatest for the 5000-, 4000-, 3000- and 2000-lb tension ranges, it is apparent that most of the fatigue damage occurred while transporting the load, rather than when picking it up.

Calculation of total cycles to failure for the various load ranges encountered is given below Table 3. Converting cycles to months, based on the estimated 125 cycles per lift and 80 lifts per day, results in an estimated chain life of 3.9 months. As discussed previously, chain failures had been experienced on this application every 3–4 months. Using the information developed, it was found that increasing the endurance limit of the same chain, by prestressing, raised the endurance limit sufficiently to give the life desired. No failures have occurred in the three years since the new chain was installed. Without the knowledge of dynamic chain loading, a much heavier chain would have been selected.

Table 4 compares the prediction of the Corten-Dolan equation with that of Miner's. In this particular application, the Corten-Dolan relationship predicts that the tension ranges between 2000 and 4000 lb did most damage. Since these tensions are below the endurance limit of the chain, they are neglected in the Miner analysis and the latter predicts about three times longer life than was obtained.

Summary

Further application of the Corten-Dolan relationship to the prediction of total cycles to failure for heterogeneous load spectrum has been made on chains operating over sprockets on testing machines in the Chain Belt Co. In all cases tested to date, this relationship has given good correlation with the experimental evidence. Thus, for random dynamic loading, the means for predicting total cycles to failure using the cumulative fatigue damage theory proposed by Corten and Dolan correlates well with the experimental evidence obtained by our laboratory. The information required to apply the theory to practice is:

1. Knowledge of the endurance limit diagram for the part or machinery in question.
2. Knowledge of the service load spectrum.
3. The cumulative fatigue damage relationship.

While it may appear that obtaining the information required for items 1 and 2 is impractical, to do so will undoubtedly be required to predict limited life accurately no matter what further developments in the theory may be forthcoming. They are the basic elements of any strength of materials problem, be it static or dynamic.

current STORABLE

THE BASIC TREND in storable liquid propellants is directed toward combining high impulse with high-density impulse in low freezing systems. The authors believe that storable liquid propellants with at least 10% higher specific impulses and 20% higher density impulses will be available in the future.

Because at present no single oxidizer or fuel exists which combines all the desired attributes of an ideal propellant, the needs of each missile system must be analyzed when choosing a fuel. For example, maximization of specific impulse is increasingly important as the missile range increases. On the other hand, density impulse and wide liquidus range are essential in short-range missiles geared to high mobility prior to launching.*

THE DISTINCTION between storable and nonstorable rocket propulsion systems is primarily a distinction between oxidizers. Except for ammonia and hydrogen, the energetically useful liquid rocket fuels such as the hydrocarbons, alcohols, amines, and hydrazines are generally stable, easily handled liquids under ambient conditions. The degree of propulsive utility which can be derived from the fuel is a direct function of the relative amount of oxygen and/or fluorine available in the oxidizer in an unreduced state. Elemental fluorine, followed by elemental oxygen, are inherently the highest performance oxidizers[1] (that is, yielding the highest propellant specific impulse) for any propellant system regardless of the chemical nature of the fuel. The cryogenic character of these oxidizers is the price nature elicits in exchange for their performance potential. All other oxidizers, solid or liquid, are inherently inferior because of the available energy loss involved in combining these oxidizing elements into a molecular structure with other elements of inferior oxidation potential. The performance of some high-energy cryogenic propellant systems is shown in Table 1 and compared with the performance of the standard cryogenic system, liquid oxygen and RP-1, the propellant combination which powers current liquid IRBM and ICBM missiles.

Table 1 — Performance of Cryogenic Propellant Systems

(1,000 psia, shifting equilibrium, sea level optimum expansion)

	WMR[a]	$\bar{d}$	$I_{\bar{d}}$	I_s
F_2/N_2H_4	2.30	1.31	476	363
F_2/H_2	7.60	0.45	185	410
O_2/N_2H_4	0.90	1.07	282	313
O_2/H_2	4.02	0.28	109	391
O_2/RP-1	2.56	1.02	306	300

[a] Weight-mixture ratio.

The question of storability resolves itself, therefore, to a compromise between propellant performance and physical properties; the extent of the compromise will vary with the mission and missile system requirements. The concept of storability may have several interpretations. The most demanding interpretation places the following requirements on the propellant system:

1. Completely sealed tanks.
2. Temperature range of – 65 to + 160 F (liquidus).
3. High density.
4. Two–five years storage.
5. No change in propellant assay or gas pressure rise.
6. Nonreactive with materials of construction.
7. Small temperature effect on important phys-

* Paper presented at SAE National Aeronautic Meeting, New York, April 3, 1959.

[1] The practicality of ozone as an oxidizer is still in doubt and hence is not considered within the context of the discussion.

trends in

PROPELLANTS

J. Silverman and S. A. Greene
Rocketdyne Division, North American Aviation, Inc.

ical properties.

During the past decade, rocket experience and chemical logic have clearly suggested that high propellant performance and strict adherence to the tabulated requirements tend to be mutually exclusive, particularly with respect to the wide liquidus temperature range. As we shall see during a subsequent review of storable propellants, the restrictive nature of the classical concepts would eliminate such high-performance propellants as hydrazine, nitrogen tetroxide, and chlorine trifluoride. This circumstance, together with the rapid growth in mechanical technology as well as special purpose missiles, has led to relaxed liquid temperature range requirements for many applications, thus permitting full exploitation of the combined attributes of storability as well as the high performance and versatility available from liquids.

Attributes of Storable Liquid Systems

Performance — The measure of thermodynamic performance of any rocket propellant is characterized by the specific impulse, I_s. For optimally expanded nozzles, the I_s is precisely proportional to the square root of the isentropic specific enthalpy change between the combustion chamber and the nozzle exit. Using ideal gas assumptions, the specific impulse can be calculated from basic thermodynamic data such as enthalpies, entropies, heat capacities, and chemical equilibrium constants, provided these data are known for the reactants and all the important combustion products. Approximate and useful formulas can be derived for specific impulse, all of which can be reduced to the general expression:

$$I_s = k\sqrt{T_c/M, f(\gamma)} \qquad (1)$$

where:

T_c = Combustion-chamber temperature
M = Average molecular weight of combustion products
$f(\gamma)$ = Function of γ, the isentropic expansion coefficient

The influence of γ on the specific impulse equation is generally minor compared to the influence of the parameter T_c/M, provided that excessively large compositional changes or condensation effects do not occur during the expansion. Hence, the relative magnitude of $\sqrt{T_c/M}$ is a useful guide to the specific impulse capabilities of a propellant. The maximation of this parameter is sought for high impulse. This is accomplished preferentially through any path which will reduce the magnitude of M rather than increasing T_c, that is, the use of nonoxidized, hydrogen-rich fuels which are, because of chemical considerations, invariably liquid in character. With standard oxygen containing oxidizers, this approach permits the attainment of high performance at moderate combustion temperatures, thus minimizing the thermal stresses on the rocket motor itself.

The ability to cool the rocket motor regenera-

tively by using the fuel as the coolant permits the use of fluorine-rich oxidizers or lightweight, high-heat producing metals, such as boron, aluminum, and the like to raise the value of T_c. Metallized fuels, however, have questionable value for several reasons: (1) They yield high molecular weight combustion products, and (2) the solid oxides produced can result in large nozzle losses due to thermal lag and hydrodynamic drag effects.

In summary, the storable liquid bipropellant offers a convenient, trouble-free path to high specific impulse by proper choice of hydrogen-rich fuels as well as oxidizers containing oxygen and fluorine in a thermochemically highly reactive state. A later discussion will illustrate the performance levels which can be attained.

Instant Readiness — Missile propulsion systems in which the propellants are sealed in the missile tanks are indeed in a full state of instantaneous readiness. The missile can be fired as soon as it can be aimed.

Hypergolicity — Many high-energy storable liquid systems have the additional advantage of an inherently simple and extremely reliable ignition system, since they are hypergolic (spontaneously combustible on fuel-oxidizer contact) and remain so even at very low pressures (high altitudes) and temperatures. Hypergolicity also confers a built-in restart capability of exceptional simplicity and reliability. A further implication of hypergolicity relates to the absence of violent destructive explosive hazards, since no large accumulation of mixed propellants can occur. Nonhypergolic systems present fire as well as explosive hazards. Fortunately, hypergolic ignition can usually be attained without sacrifice of other desirable propellant attributes.

Growth Potential and Versatility — Liquid systems have enormous growth potential and inherently resist obsolescence. From the propellant viewpoint, it is frequently feasible to substitute new, higher performing propellants for existing operational propellants as they become available from the chemist's laboratory. An illustrative example is the substitution of Hydyne for alcohol in the Redstone engine to yield the 10% increase in engine performance required to throw the Explorer satellites into orbit. It is of interest to note that in this case the time interval from the initiation of desk-top analytical studies through and including large engine acceptance testing was only six months. Similarly, interconversion between storable and nonstorable propellant systems is equally practical. It is not necessary to build expensive new engines to incorporate chemical advances. Storable liquid engines can also be scaled in thrust over a considerable range by adjustment of flows and pressures without major redesign. Duration can be varied at will virtually without limit by changing tankage volumes.

Mobility — The physical separability of the fuel and oxidizer from the missile makes for easier mobility. Propellants can be transported to the launching site by the most convenient, most economical commercially appropriate method; the missile, relieved of the huge mass of propellant, reverts to light, easily transportable hardware. Of course, since the propellants are storable, the tactician has full freedom of choice and may act as the occasion demands; the missile can be handled charged or empty.

Low Maintenance Costs — The ability to maintain a missile in a fully tanked and ready condition also reflects in reduced equipment and manpower requirements. The extent of the cost reduction over cryogenic systems depends in some measure on the nature of the system.

Table 2 — Physical Properties of Selected Oxidizers

(at 68 F)

	WFNA	IRFNA	N_2O_4	MON 70/30	95% H_2O_2	CIF_3
Boiling Point, F	190	150	70.1	− 4	294	53.2
Freezing Point, F	−52	−57	11.8	−112	21.9	−105.4
Density, gm/ml	1.50	1.57	1.45	1.37	1.42	1.83
Viscosity, cs	0.54	0.89	0.31	—	0.885	0.238
Vapor Pressure, psia	0.83	2.5	14.0	87	0.032	20.57
Heat Capacity, Btu/lb-F	0.42	0.39	0.75	—	0.643	0.40

Table 3 — Performance of Mixed Oxides of Nitrogen (MON) with Hydyne

(500 psia, frozen equilibrium, sea level optimum expansion)

% NO in N_2O_4	WMR	d	I_d	I_s
0	2.57	1.22	321	263
10	2.58	1.20	314	264
20	2.63	1.19	315	265
30	2.68	1.18	314	266

Table 4 — Physical Properties of Selected Fuels

(at 68 F)

	RP-1 JP-4,5	HYDYNE	UDMH	N_2H_4
Boiling Point, F	350–500	140–400	146	236
Freezing Point, F	−58	−65	−71	35
Density, g/ml	0.81	0.86	0.79	1.01
Viscosity, cs	2.5	1.45	0.71	0.96
Vapor Pressure, psia	0	2.49	2.3	0.20
Heat Capacity, Btu/lb-F	0.45	0.65	0.65	0.74

Storable Oxidizers

Nitric Acids — Modified nitric acids have been the most widely used storable oxidizers heretofore. Pure anhydrous nitric acid (WFNA) is an oxidizer which cannot be used in classical storable applications. The material is not suitable for two reasons, the first being that the corrosion of aluminum and steel alloys is appreciable, thus seriously limiting the acceptable storage life. The container material will go into solution, the composition of the acid is variable, and eventually a solid phase is usually precipitated. Corrosion of certain aluminum and stainless-steel alloys has been essentiallly eliminated by the addition of 0.5% by weight HF, which passivates container walls by the deposition of a metallic fluoride coating. Acid which contains HF is usually referred to as an "inhibited" nitric acid. However, even "inhibited" acid cannot be stored indefinitely in many steel and other alloy tanks which are desirable from the standpoint of high strength-to-weight ratio. And secondly —

Pure nitric acid is thermodynamically unstable and slowly decomposes to yield oxygen, nitrogen dioxide, and water. In sealed storage, an excessive equilibrium pressure primarily due to the oxygen component is theoretically predicted and has been experimentally confirmed. This pressure reaches about 1100 psia at 130 F. Thus, operational acid mixtures must contain approximately 3% water and 15% by weight of nitrogen dioxide in order to suppress high oxygen equilibrium partial pressure by a mass action effect. Such a mixture will exhibit an equilibrium pressure of no more than 30 psia at 130 F. Physical properties of the acids and other oxidizers are listed in Table 2. IRFNA (inhibited red furming nitric acid—the red color is imparted by nitrogen dioxide) is the standard oxidizer utilized in currently operational liquid propellant air-to-air, air-to-surface, and surface-to-air missiles. Advantages are extremely low cost, excellent availability, high density, and good physical properties over extended temperature ranges. The ground storability requirements are easily met. The main disadvantage is the modest performance delivered by the oxidizer with fuels of interest. Well-known applications of nitric acid include, for example, the second stage of the Vanguard and Thor-Able, the Army Corporal, the Aerobee, and the Nike-Ajax sustainer rocket.

Oxides of Nitrogen — Nitrogen dioxide offers performance levels above those delivered by nitric acids. A further advantage is that if the water content is kept below 0.1% (the material is usually manufactured to this specification) a wide variety of metallic materials including high-strength mild steel can be used as a material of construction. Nitrogen tetroxide suffers from the disadvantage of having an extremely narrow liquidus range, and will not meet the ground storability requirements of – 60 to + 160 F. However, the low cost, good density, easy availability, good materials compatibility, and high performance favor this oxidizer for systems that do not require the lowest freezing points. The relatively low boiling point is not an important disadvantage as was previously discussed.

An additive which can depress the freezing point of N_2O_4 is the low boiling gas, nitric oxide. The mechanism of depression is believed to involve the formation N_2O_3 which is soluble in nitrogen tetroxide. In addition, the specific impulse of propellant systems which utilize NO is seen to rise slightly, as is shown in Table 3, while the density impulse is decreased.

Addition of 25% by weight of NO to N_2O_4 reduces the freezing point to – 65 F while the vapor pressure rises rather sharply. As is shown in Table 2, the vapor pressure at 68 F rises from 14 to 87 psia and is approximately 500 psia at 160 F.

Hydrogen Peroxide — Hydrogen peroxide has received a rather large amount of attention, although it cannot meet the previously stated requirements for a storable oxidizer. As can be seen from Table 2, the high freezing point severely limits its operational temperature range. More serious is the fact that peroxide decomposes slowly during storage, yielding water and oxygen, which necessitates periodic venting of tanks. Peroxide is satisfactorily storable in missile engines for periods of a few days to perhaps a few weeks. Peroxide can be satisfactorily stored in large vented tanks for a year or more with negligible decomposition.

Chlorine Trifluoride — This unusually dense and most vigorous oxidizer is a leading candidate for future exploitation. The density is almost 30% greater than that of conventional oxygenated oxidizers. Its incendiary nature even with water insures reliable hypergolic ignition with all fuels of interest. The low freezing point is advantageous whereas its boiling point is sufficiently high for practical utility for almost all applications. Although the reactivity of chlorine trifluoride would suggest severe problems in terms of compatibility with materials of construction, standard as well as new techniques of passivation have obviated difficulties. There are available an adequate number of materials of construction including metals, plastics, and elastomers.

Chlorine trifluoride is a very high performing storable oxidizer delivering specific impulse that is essentially identical with nitrogen tetroxide with fuels of interest. The overall propellant bulk density at the same time is increased by approximately 16%. A disadvantage of the oxidizer is that systems operate at somewhat higher oxidizer-to-fuel mixture ratios than the corresponding oxygenated systems, and hence, less fuel is available for regenerative cooling. In addition, chamber temperatures are about 1000 F higher than with oxygen-based systems exposing the coolant fuels to large thermal stresses. However, chlorine trifluoride may also prove suitable for use as a coolant, since it is stable and has fairly good physical properties. The presence of large quantities of HF in the exhaust may be disadvantageous in some applications. Actually, this is a greater problem in static engine testing during development than it will be in operation, where the rapidly accelerated flight of the missile will assure harmless dispersion of the exhaust gas.

Storable Fuels

Hydrocarbons — All large ballistic missile rocket engines now use hydrocarbon-type fuels. The amount of hydrocarbon fuels fired in rocket engines is larger by many orders of magnitude than any other fuel. The principal ballistic missile hydrocarbon fuel is RP-1 which resembles the kerosene jet fuels used in air-breathing engines, but is of a much narrower cut and contains little or no aromatics and olefins, thus considerably improving the stability and storage properties. Typical of petroleum refinery products, hydrocarbon fuels are cheaper by a factor of 10 as compared to other essentially pure chemical, storable fuels. Physical properties of RP-1, as well as some other storable fuels, are listed in Table 4.

Densities of hydrocarbon fuels are rather low, but the overall propellant bulk density is generally good due to the high mixture ratios employed for maximum specific impulse. Another consequence of the high mixture ratio is that relatively little fuel is available for regenerative cooling and one may be forced to employ oxidizer cooling.

Other disadvantages of using hydrocarbon fuels include modest specific impulse, severe combustion stability problems in some cases when used with nitric acid, and the absence of reliable hypergolicity

with storable oxygenated oxidizers. Hydrocarbon fuels are generally unsuitable for use with fluorine based oxidizers.

Amines and Hydrazines — The most important storable fuels are conveniently classified into categories which reflect the number of nitrogen atoms in the molecules in the case of monoamines and polyamines, while hydrazine molecules contain the azine (N-N) linkage. The most important hydrazine-type fuels in the order of decreasing performance are hydrazine, monomethyl hydrazine (MMH), and unsymmetrical dimethyl hydrazine (UDMH).

The amines result in only moderately higher performance over hydrocarbons (Table 5). A broad spectrum of physical properties (as shown in Table 6), coupled with selective mutual miscibility, permits the tailoring of optimized fuel formulations for every mission. Advantage of the mono- and polyamines include reasonable cost and stable combustion with the principal storage oxidizers. The potential of monoamines has not yet been realized, but monoamines offer attractive features for some special purpose missiles. Low cost of fuel is coupled with optimum performance at relatively high mixture ratios, suggesting the feasibility of oxidizer cooling.

Application of polyamines to storable propellants has taken place by utilizing mixtures with the high performing, but more costly, member of the hydrazine family, UDMH, resulting in a significant fuel price reduction. Operational mixtures of diethylenetriamine and UDMH, although yielding a specific impulse some 1% lower than with UDMH alone, realize a 3% advantage in density impulse with storable oxidizers. High molecular weight polyamines suffer from rather excessive viscosities at low temperatures. Aside from cost, UDMH and MMH would appear to be the most favored of storable fuels where low-temperature requirmement must be accommodated. The physical properties are excellent, performance is fairly high, and stringent freezing point requirements are easily met.

The conventional fuel which offers the highest density and specific impulse available today, no matter which oxidizer is utilized, is hydrazine (Table 7). Its physical properties are generally good, but the high freezing point of hydrazine has precluded its use for many storable applications. Freezing point depressants for hydrazine have long been sought, but none have been found which do not excessively compromise the performance or stability. The development of hydrazine as a storable propellant appears rather promising for the larger ballistic missiles, since the environmental launch condition can be controlled to accommodate the high freezing point of hydrazine. A further advantage of hydrazine lies in that optimum performance is obtained at mixture ratios which are a minimum, resulting in large amounts of fuel for regenerative cooling. The fuel is an excellent coolant by virtue of its physical properties.

Table 5 — Performance of Amines with IRFNA

(1,000 psia, shifting equilibrium, sea level optimum expansion)

	WMR	d	I_d	I_s
Monoamines				
dimethylamine	3.88	1.22	330.3	270.7
triethylamine	4.28	1.29	348.1	269.9
Polyamines				
ethylenediamine	3.09	1.33	361.1	271.5
diethylenetriamine	3.26	1.37	368.9	269.3
Hydrazines				
hydrazine	1.47	1.28	362.4	283.1
MMH	2.47	1.28	357.1	279.0
UDMH	2.99	1.26	348.3	276.4

Storable Propellant Performance

The propellant combination WFNA/JP-4 and later IRFNA/JP-4 were the first storable systems given serious consideration. Performance of IRFNA with fuels of interest is shown in Table 8. In addition to the previously mentioned difficulties with WFNA, problems which caused the abandoning of these propellants were the absence of reliable hypergolic ignition and unstable combustion. The addition of approximately 40% of UDMH to JP-4 solved both the combustion and ignition difficulties with IRFNA. The amine, UDMH is reliably hypergolic with IRFNA, and also increases the performance by 2%.

A further increase in performance, while maintaining a completely storable system, was realized by the elimination of the hydrocarbon portion of the fuel by substitution of another dense amine fuel, diethylenetriamine. Performance was increased by 4% over the original hydrocarbon fuel while the physical properties and storability were also considerably improved. The fuel mixture 60% UDMH-40% diethylenetriamine is known as Hydyne and is used in the Redstone missiles which currently launch the Explorer satellites.

Although the system IRFNA-Hydyne has good storability, wide operational temperature limits, is readily available, cheap, and hypergolic, the performance is relatively low. Substitution of UDMH for Hydyne increases the specific impulse by about 1%, but the overall performance, which also depends on the propellant bulk density will not be im-

Table 6 — Physical Properties of Some Amines and Hydrazines

(at 68 F)

	Monoamines		Polyamines		Hydrazines		
	dimethyl amine	triethyl amine	ethylene diamine	diethylene triamine	N_2H_4	MMH	UDMH
Boiling Point, F	44	193	242	405	236	192.5	146
Freezing Point, F	− 134	− 155	50	− 38	35	− 62.5	− 71
Density, g/ml	0.66	0.73	0.90	0.95	1.01	0.878	0.79
Viscosity, cs	0.29	0.49	1.8	8.0	0.96	0.99	0.71
Vapor Pressure, psia	25	1.03	< 0.1	< 0.1	0.20	0.75	2.3
Heat Capacity Btu/lb-F	0.73	0.52	0.70	0.65	0.74	0.70	0.65

Table 7 — Performance of Oxidizers with N_2H_4

(1,000 psia, shifting equilibrium, sea level optimum expansion)

Oxidizer	WMR	$\bar{d}$	I_d	I_s
IRFNA	1.47	1.28	362	283
95% H_2O_2	2.17	1.26	355	282
N_2O_4	1.34	1.22	356	292
ClF_3	2.77	1.51	444	294

Table 8 — Performance of Fuels with IRFNA

(1,000 psia, shifting equilibrium, sea level optimum expansion)

Fuel	WMR	$\bar{d}$	I_d	I_s
RP-1	4.80	1.35	362	268
Hydyne	3.11	1.31	358	273
UDMH	2.99	1.26	348	276
50% UDMH-50% N_2H_4	2.20	1.27	354	279
N_2H_4	1.47	1.28	362	283

Table 9 — Performance of Fuels with 95% H_2O_2

(1,000 psia, shifting equilibrium, sea level optimum expansion)

Fuel	WMR	$\bar{d}$	I_d	I_s
RP-1	7.35	1.30	355	273
Hydyne	4.68	1.27	351	276
UDMH	4.54	1.24	345	278
50% UDMH-50% N_2H_4	3.35	1.25	349	279
N_2H_4	2.17	1.26	355	282

Table 10 — Performance of Fuels with N_2O_4

(1,000 psia, shifting equilibrium, sea level optimum expansion)

Fuel	WMR	$\bar{d}$	I_d	I_s
RP-1	4.04	1.25	345	276
Hydyne	2.71	1.22	344	282
UDMH	2.61	1.18	336	285
50% UDMH-50% N_2H_4	2.00	1.21	348	288
N_2H_4	1.34	1.22	356	292

Table 11 — Performance of Fuels with Chlorine Trifluoride

(1,000 psia, shifting equilibrium, sea level optimum expansion)

Fuel	WMR	$\bar{d}$	I_d	I_s
RP-1	3.20	1.41	364	258
Hydyne	2.98	1.43	395	276
UDMH	3.03	1.38	386	280
50% UDMH-50% N_2H_4	2.89	1.45	416	287
N_2H_4	2.77	1.51	444	294

proved. If a missile is tanked with two sets of propellants, with identical I_s, the more dense propellant combination will obviously carry the missile to a greater range. Thus, for missiles which must, of necessity, be limited in volume, the density impulse is a significant index of performance. The density impulse obtained with Hydyne is 3% greater than that with UDMH. Although not considered a fully storable propellant, the performance of 95% hydrogen peroxide with storable fuels is shown in Table 9. Performance levels are quite comparable with IRFNA. This combination has limited storability. Hydrogen peroxide is the only even moderately storable oxidizer available today which has good performance characteristics with the cheap, readily available jet fuels. For this reason, it merits consideration when only short storage times are required.

Proceeding to an oxidizer which is more energetic than IRFNA, substitution of N_2O_4 for IRFNA will increase the specific impulse of fuels by about 3.2%, but the density impulse decreases from 2–4%, as is shown in Table 10. The oxidizer N_2O_4 does not satisfy the most stringent storability condition because of its high freezing point, but is operational with some temperature conditioning and can be made storable by the incorporation of NO as was previously shown. Note that although propellant systems utilizing mixed oxides are relatively high performing, cheap, compatible with mild steel and hypergolic, we have a rather high oxidizer vapor pressure and the density impulse is decreased.

Proceeding to the next oxidizer of interest, the fluorinated oxidizer chlorine trifluoride is quite attractive. From Tables 10 and 11 it can be seen that although the specific impulse obtained with chlorine trifluoride is somewhat less (1–2%) than that obtained with N_2O_4 (except in the case of hydrazine), the density impulse is about 16% greater for all fuels of interest. Chlorine trifluoride is thus an excellent oxidizer in terms of performance, density, hypergolicity, and low freezing point, when used with amine or hydrazine based fuels. It is excessively low in performance with hydrocarbons.

Summary

It is apparent that no single oxidizer or fuel exists which can be designated as an ideal or "best" propellant, that is, combining all the desirable attributes without compromise. No single propellant system can satisfy all the requirements for all conceivable missile systems. Each missile system dictates its own unique set of required propellant properties as well as the order of importance of these properties. For example, maximization of specific impulse becomes increasingly important as the range of the missile increases whereas, concurrently, less importance is placed on properties such as density and low freezing point. On the other hand, density impulse and wide liquidus range become key factors in short-range missiles geared to high mobility prior to launching.

It is, nevertheless, true that the basic trend in storable liquid propellants is actively in the direction of combining high impulse with high-density impulse in low freezing systems. Highly directional research is currently in progress in the fields of chemical synthesis of new oxidizers and fuels, the formulation of "tailored" mixtures for very specific applications, the development of new metallic and nonmetallic materials for use with corrosive liquids as well as the development of new passivation techniques. Although the nature of propellant research is highly classified, progress to date clearly indicates that the future generation of storable liquid propellants can be expected to have specific impulses at least 10% higher, and density impulses at least 20% higher than currently available systems. These propellants should prove advantageous in every respect over all but the very most energetic cryogenic systems.

Radiotracers

during detergent oil

Harry Halliwell
U. S. Naval Engineering Experiment Station

THE DETERGENCY level of early diesel-engine lubricating oils was found to be inadequate for the operation of submarine main propulsion engines under severe-duty operations with minimum quality fuels. With the formulation of higher detergency level oils which were developed to assure better operation of engines under these conditions, it was considered that small abrasive particles might be dispersed by the detergent additive and carried in suspension in the lube oil system, thereby imposing increased wear rates on rubbing and sliding surfaces in the engine. In addition, it was considered that the increased dispersant ability of these oils might prevent agglomeration and reduce the ability of the filters to remove these wear-promoting particles.

Consequently, a program was initiated to investigate the effect of the use of filters and the level of oil detergency of current and experimental marine diesel lubricating oils on the wear of full-scale engine components. The use of radioactive tracer techniques was considered an excellent means of investigating these parameters quickly and accurately. Since piston rings are generally the fastest wearing parts in a diesel engine, this technique was used to determine the effect on piston-ring wear of filtration, oil detergency level, and abrasive particle size.

Under normal operating conditions the rate of piston-ring wear is a measure of both the quality of

THIS PAPER describes how radiotracers were used to determine the effect of certain oil variables on wear in marine diesel engines. The tests were made with a clean engine oil system and with abrasive contaminant added, to show how filtration, oil detergency level, and abrasive particle size affected piston-ring wear.

The author concludes that detergent-type oils used in a relatively clean engine will effectively minimize piston-ring wear. Filters are necessary to protect diesel engines from gross quantities of wear debris produced as a result of dirty engine operating environments. For operations that require protection of moving parts against abrasive materials, the use of full-flow filters in the lub oil system is recommended. When essentially clean conditions prevail and when using low detergent level oils, bypass filters might be more beneficial in controlling wear rates.*

Reveal Engine Wear

filtration study

the rings and effectiveness of the lubricant. Since the physical, chemical, and metallurgical properties of all of the rings used in this program were the same, the relative effects of the different lubricants could be evaluated. In addition, by using only one lube oil base stock, different concentrations of detergent additives, filters, and other materials such as abrasive contaminants were evaluated with regard to their particular effect on engine wear in the ring-cylinder liner area.

This investigation was divided into two parts. The first part covered investigations of the effects of the oil detergency level and filtration on piston-ring wear during clean sump operating conditions, and the second part covered the effects of these factors plus the addition of contaminants and subsequent filtration on piston-ring wear in the GM 3-71 diesel engine.

Method of Test

The tests were conducted in a General Motors 3-71 series diesel engine coupled to a 150-hp electric dynamometer. The lube oil system was modified to permit the determination of piston-ring wear by use of the radioactive isotope tracer technique. A schematic drawing of the oil circulation system shows these minor modifications (Fig. 1).

The walls (A) of the crankcase pan were altered to provide steep slopes which converged to a plugged 1½ in. pipe sleeve (B). In this manner the radioactive wear debris was washed down the slopes to the base of the pipe sleeve by the flowing oil. To minimize sedimentation and maximize circulation of the wear particles, the suction end of the lube oil pump (C) was located ¼ in. from the bottom of the sleeve. Oil leaving the pump could be directed through filters (D) or (E), which were installed in the lube oil system between the pump and the main lube oil cooler (F). The full-flow filter (D) passed the total flow (approximately 10 gpm) through its element. When the bypass filter (E) system was used, a small portion (0.4 gpm) of the lube was filtered and then returned to the engine oil sump; the balance passed through the cooler (F).

The main lube oil cooler (F) was required to maintain the desired lube oil operating temperature in the engine. A needle valve (G) located in the downstream side of the cooler controlled the lube oil flow rate to the counter well (J) at 0.5 ± 0.1 gpm during each test run. The remainder was fed to the oil distributing gallery in the cylinder block. The oil flowing to the counter well was first passed through a heat exchanger (H) to control the tem-

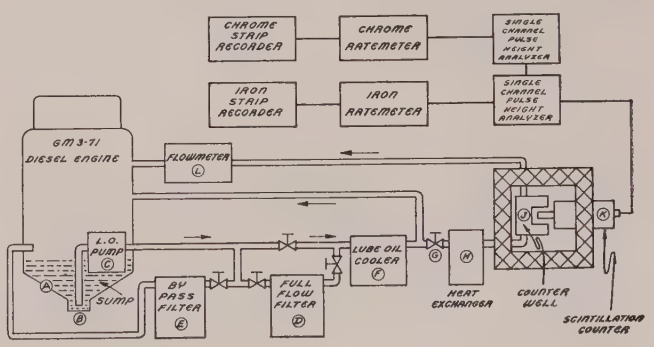


Fig. 1 — Schematic drawing of oil system

* Paper presented at SAE Summer Meeting, Atlantic City, June 16, 1959.

perature in the counter well at 100 ± 10 F. This was necessary since it was found that the photomultiplier tube in the scintillation counter (K) was sensitive to temperatures exceeding 135 F and would produce erroneous wear rates. After the lube oil passed through the lead-shielded counter well, it was returned through a flowmeter (L) to the engine oil sump.

Instrumentation

The radioactivity detector selected for this program was a Nuclear-Chicago DS-5 scintillation probe containing a 2 in. diameter × 2 in. high thallium-activated sodium iodide crystal. The face end of the crystal was butted against the inner portion of an aluminum counting well through which the oil passed from bottom to top. The probe and counting well were mounted in a lead cube with walls 4 in. thick to provide shielding from background radiation. The volume of the counting well was 560 milliliters. The signals from the scintillation crystal were received in the amplifier section of one of the two Nuclear-Chicago model 1812 single channel, pulse height analyzers. The amplified signals were then sent to the discriminator sections of each pulse height analyzer where the activity of the Fe^{59} and Cr^{51} isotopes was measured simultaneously. The output of each single channel analyzer was indicated on a Nuclear-Chicago Model 1620 ratemeter, set at a 50-sec integrating time constant to reduce fluctuations, and then recorded on an Esterline-Angus strip chart recorder. In this manner chrome (ring face to cylinder liner) and iron (ring back and sides to piston groove) wear were measured.

Piston Rings

Chrome-plated piston rings having identical physical, chemical, and metallurgical properties were activated at the Brookhaven National Laboratory in a neutron flux of 3×10^{12} neutrons per square centimeter per second. Ten days in the pile were necessary to obtain the desired activity. Nine rings, placed in the second, third, and fourth grooves of each piston, gave a total activity of 1665 millicuries at time of arrival on the Station. Since it was considered that more representative wear due to lube oils would occur on the three lower compression rings, the top ring on each piston was not activated. Chrome-faced rings were used because they are the standard Navy part for the GM 71 series diesel engines and because their use would permit the discrimination between face wear and side plus back wear of the rings.

Test Filters

New full-flow and bypass filter elements of the multifold paper type recommended for service with the GM 71 series engines were used for each test during the filtering effectiveness studies.

The full-flow and bypass filters consisted of a replaceable element enclosed with a case that was mounted on the side of the engine. A bypass valve, which opens at 15 psi, was built into the full-flow adaptor to insure adequate engine lubrication even if the filter were to become plugged. The full-flow filter elements are rated by the manufacturer to remove the larger foreign particles, 40 microns or more, that may be present in the oil, without unduly restricting the flow of oil. During the full-flow filter tests, all of the lube oil supplied to the engine, approximately 10 gpm, passed through the filter element. During the bypass filter tests, a smaller portion of the lube oil was filtered, approximately 0.4 gpm. This filtered oil was returned directly to the engine crankcase. The bypass elements are designed to remove minute foreign particles, 5 microns or more, that may be present in the oil.

Table 1 — Piston-Ring Wear (Chrome) Detergency and Filter Results

Oil Code	Chrome Wear Rate, 10^{-7} gm/hr, Filter System			
	None	Full-Flow	Bypass	Average
A	260	249	158	222
B	225	204	196	208
C	142	112	118	124
D	42	34	35	37
Average	167	150	127	

Table 2 — Piston-Ring Wear (Iron) Detergency and Filter Results

Oil Code	Iron Wear Rate, 10^{-7} gm/hr, Filter System			
	None	Full-Flow	Bypass	Average
A	98	95	80	91
B	85	82	65	77
C	63	57	58	59
D	30	27	28	28
Average	69	65	58	

Contaminant Materials

The contaminants used in these tests were natural aluminum oxide abrasive compounds with average particle size grades of 5, 15, and 22 microns in diameter. For the abrasive partice size studies the aluminum oxide was introduced into the crankcase during each test in the amount of approximately 10 parts per million of the total lube oil charge (by weight).

Procedures

Engine Test Conditions — Standard engine operating conditions employed for all test runs were as follows: 67.5-bhp load, 1800-rpm speed, 200 F lube oil temperature, 170 F water jacket temperature, and an oil flow rate of 0.5 ± 0.1 gpm through the scintillation counter well.

Clean Sump Operation (Part 1) — Each test was conducted as follows: The engine was charged with exactly 5 gal of new test oil. The engine was started and allowed to warm up for approximately 15 min at low speed and load until the lube oil temperature reached 200 F. Then the speed and load were increased to 1800 rpm and 67.5 bhp, respectively. Each test was run for 6 hr, and the increase in radioactivity level of the oil during the test period was recorded. Equilibrium wear rates were reached, usually, within 1 hr. The data recorded during the last 5 hr of the test were used to establish the wear

Table 3 — Detergent Additive Level Effect on Wear

Oil	% Detergent (by volume)	Relative Wear, %		Total
		Chrome	Iron	
A	0	100	100	100
B	2	94	85	89
C	5	56	65	60
D	16	17	31	24

Table 4 — Effect of Filtration on Wear

Oil Code	Wear, % of that with No Filter	
	Full-Flow	Bypass
A	96	67
B	92	84
C	82	86
D	85	88

Table 5 — Chrome Versus Iron Wear

Average wear rates, 10^{-7} gm/hr

Type Wear	None	Full-Flow	Bypass	% Total Wear
Chrome	167	150	127	70
Iron	69	65	58	30

rate for each particular test. At the end of each test, the used oil was drained immediately and a flush oil was used for 15 min to remove residual wear debris from the lube oil and counting systems. To minimize recording errors, an instrument calibration check, using a standard solution containing irradiated chromium and iron ions, was made at frequent intervals throughout the test program.

A randomized block of the Latin Square design was used to establish the order of the test runs during this phase of the investigation, in an attempt to fairly determine differences between the oil detergency levels and the type of filtration.

Contaminated Sump Operation (Part 2) — This procedure was designed to evaluate the effects on wear of filtration and abrasive particle size in oils of different detergent additive levels. Natural aluminum oxide abrasive compounds, with average particle sizes of 5, 15, or 22 microns in diameter, were used as the lube oil contaminant. Two lube oils, derived from the same base stocks and formulated by the same refiner, were used. These oils contained the same detergent-type additives in amounts of approximately 2 and 16% by volume. The stock MS 1065, containing no detergent additive, was used to determine base line engine wear data. Each run consisted of 18 hr at the previously listed standard engine operating conditions. The first 6 hr established a constant wear rate for the noncontaminated test oil being used. Then 10 ppm of one grade of the abrasive compound was mixed in a homogenizer with a small portion of the test oil. This material was added to the crankcase at the end of the sixth hour and the new wear rate was recorded for another 6 hr. At the end of the twelfth hour, a new full-flow or bypass filter element was incorporated into the lube oil system. In this manner, it was possible to evaluate the effects of detergent additive, abrasive particle size, and filtration on chrome (face) and iron (sides and back) wear of the piston rings.

Finally, an attempt was made to determine if airborne dust would have a supplementary effect on the wear rates developed during these tests. Bon-Ami scouring powder was passed through a 200-mesh wire screen into the air cleaner of the engine at a rate of approximately 10 gm/hr for a 5-hr period. No noticeable increase in the chrome or iron wear rates was detected.

Results of Tests

Detergency and Filter Evaluation (Part 1) — A tabulation of the results of the detergency and filtration tests are shown on Tables 1 and 2.

During the period of these tests, one replacement of all the radioactive piston rings was necessary, because their activity level had decayed with attendant loss in instrument detection sensitivity. All the data in the tables to follow are presented as the amount of the weight loss of the radioactive piston rings per hour of engine operation. The computations included the factors for the decay of the activity level of the rings, so that the results are directly comparable regardless of the activity of the rings at the time of the test.

Detergent Additive Effectiveness — The average chromium and iron wear rate values for the oils containing different amounts of detergent additive indicate that wear decreases with an increase in the amount of additive blended in the oil. Table 3 shows the relative effect of the detergent additive on both chrome and iron piston-ring wear. The nonadditive oil, Code A, was used as a baseline.

Effect of Filters on Wear — The relative effects of full-flow and bypass filters in removing the radioactive wear debris from the lube oil system are shown in Table 4. These values demonstrate the ability of the filters to remove engine-produced wear contaminants in oils containing varying amounts of detergent. Data collected during operation without filters were used as the base line.

Based upon the amounts of wear debris generated with each of the four oils, averaging chrome and iron wear, the bypass filter is superior with oil Codes A and B. With two high-level detergency oils, the two types of filter media are approximately equal in effectiveness.

These results indicate a trend that with increasing amounts of detergent, the relative effectiveness of the full flow-filters improves and that of the bypass filters decreases.

Piston-Ring Face Wear Versus Ring Sides and Back Wear — Since the chrome and iron wear rates were measured individually by means of pulse height analyzers, it was possible to determine the wear of the ring faces and the ring sides plus back, independently. Table 5 shows that the face wear (chrome) contributes approximately 70% of the total wear of the piston rings. It is also of interest that the wear of all sides of the rings is reduced as the oil detergency level is increased.

Repeatability of Results — The overall repeatability of this series of tests (Part 1) was determined by dividing the mean of each group of replicate results by its standard deviation and then cal-

culating the average of these values for all tests. By this method the repeatability of the tests was found to be 22.4%.

Abrasive Wear and Filtration Results (Part 2) — An abrasive contaminant was intentionally added to the crankcase oil to simulate dirty engine operating conditions immediately after the first 6 hr of baseline engine operation. A new wear rate, showing the effect of the abrasive, was recorded for another 6-hr period. Then either a full-flow or a bypass filter was cut into the lube oil system to determine the effectiveness of the filters in removing the wear debris generated by the contaminant and indirectly, its effectiveness in removing the contaminant. Rates of accumulation of wear debris for each period (baseline, abrasive and filter) are shown graphically on Figs. 2–7, and tabulated in Tables 6 and 7.

Effect of Abrasive particle Size on Wear — In order to establish the true effect of abrasive particle size on piston-ring wear, it was necessary to determine the difference between the baseline wear and the wear rate produced after the contaminant was added to the crankcase.

The average wear rates (both chrome and iron) of all three oils for each abrasive are

(total wear, 10^{-7} gm/hr):
22-micron particle size — 1542
15-micron particle size — 2302
5-micron particle size — 1830

It will be noted that the 15-micron particle size abrasive produced considerably more wear than the 5 and 22 micron sizes with all three oils. Similarly, the average wear rates (both chrome and iron) for each oil are shown in Table 8. The two detergent additive oils (Codes B and D) apparently disperse all three particle sizes to the extent that wear with these oils was approximately two times greater than with the no-additive (Code A) oil.

Filter Effects — The effect of filtration was determined for each test as being the difference between the rate of increase of wear debris in the oil after abrasive addition and the rate of change of wear debris during the filtration period. Referring to Figs. 2–7, this means comparing the slope of the debris accumulation curve for the period 6–12 hr with the corresponding slope for the period 12–18 hr. Table 9 shows tabulations based on this com-

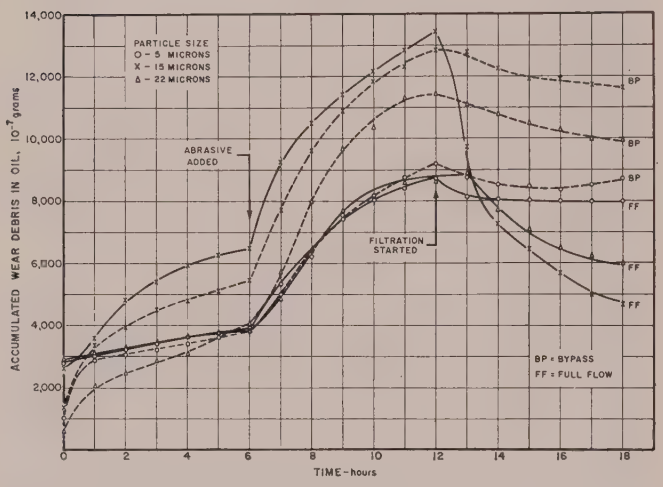


Fig. 2 — Chrome wear debris in Code A oil

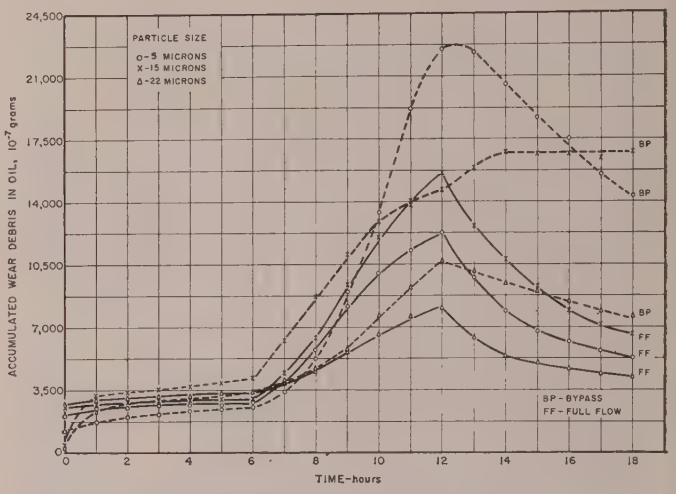


Fig. 4 — Chrome wear debris in Code D oil

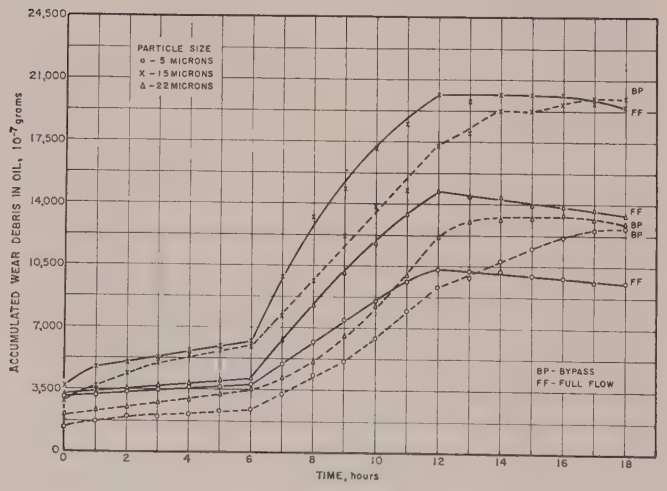


Fig. 3 — Chrome wear debris in Code B oil

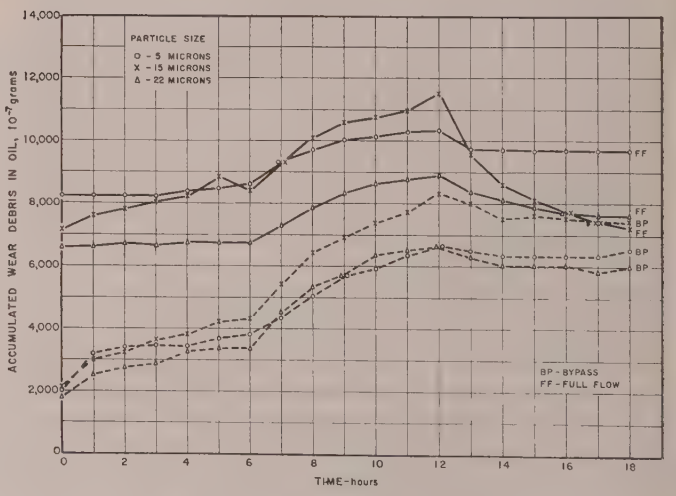


Fig. 5 — Iron wear debris in Code A oil

parison in which rate of debris removal is expressed as a percent of the rate of generation of debris. (A constant debris generation rate is assumed for the filtration period, equal to the rate for the immediately previous period. A 100% removal rate would result in a zero slope of the wear debris accumulation curve during the filtration period, since the amount generated would equal the amount removed. A negative slope to the curve would represent a greater than 100% removal rate.)

In similar manner, Table 10 shows the effect of the oil detergency level on filtration, averaged for all abrasive particle sizes. These results show that both types of filters were effective in removing from the oil system large amounts of wear debris generated by the abrasive contaminants. The differences in filtration effectiveness among the three abrasive particle size groups are not considered to be significant. However, the full-flow filters were superior to bypass filters in removing the wear producing contaminants.

From Table 10 it might appear that using filters with Code B oil is less effective than with oil Codes

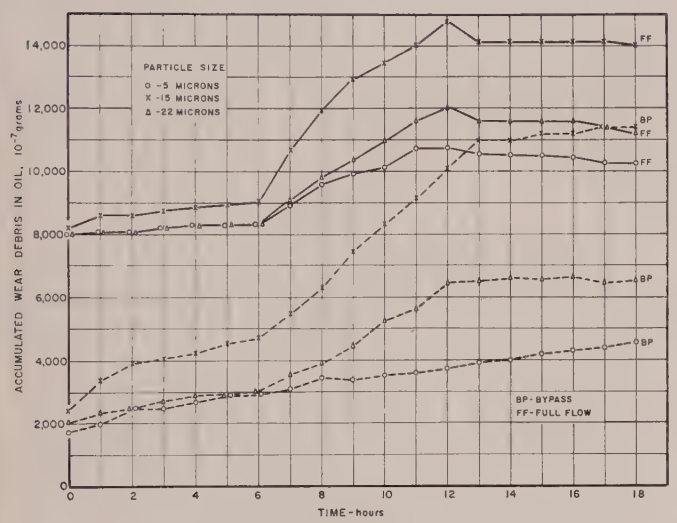


Fig. 6 — Iron wear debris in Code B oil

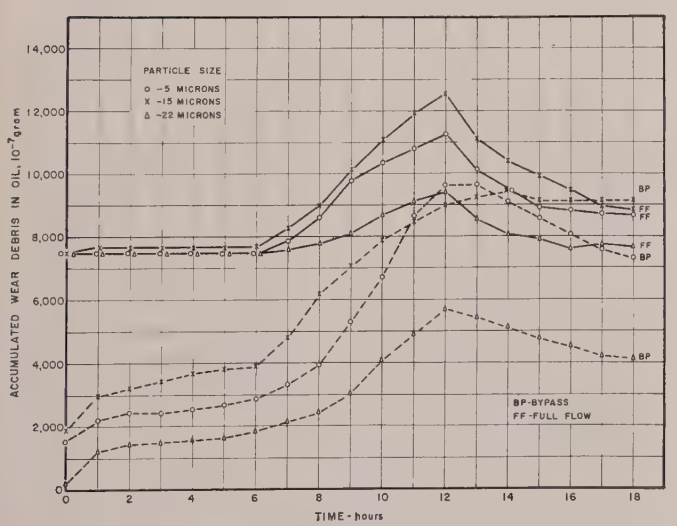


Fig. 7 — Iron wear debris in Code D oil

Table 6 — Rate of Accumulation of Wear Particles in Oil, 10^{-7} gm/hr

A. Baseline Wear — No Abrasive, No Filter

Oil Code	Chromium Wear			Iron Wear		
	1	2	3	1	2	3
A	186	295	200	38	304	53
B	128	222	145	117	193	88
D	96	68	91	0	53	79

B. Abrasive Added, No Filter

Oil Code	Average Particle Size, Microns					
	22	15	5	22	15	5
A	803	1161	812	534	753	435
B	1785	2299	1108	940	1386	534
D	807	2432	1572	505	1211	916
Average	1132	1964	1164	660	1117	628

C. Full-Flow Filter in Circuit

Oil Code	Average Particle Size, Microns					
	22	15	5	22	15	5
A	−583	−1456	− 123	−111	−348	− 50
B	−220	− 233	− 127	− 93	− 92	− 50
D	−656	−1497	−1162	−149	−319	−233

Table 7 — Rate of Accumulation of Wear Particles in Oil, 10^{-7} gm/hr

A. Baseline Wear — No Abrasive, No Filter

Oil Code	Chromium Wear			Iron Wear		
	1	2	3	1	2	3
A	389	385	185	170	240	130
B	220	350	115	120	210	202
D	200	200	118	118	180	142

B. Abrasive Added, No Filter

Oil Code	Average Particle Size, Microns					
	22	15	5	22	15	5
A	1236	1228	898	540	663	470
B	1457	1865	1153	570	902	145
D	1208	1755	3333	647	860	1149
Average	1300	1616	1795	586	808	588

C. Bypass Filter in Circuit

Oil Code	Average Particle, Size, Microns					
	22	15	5	22	15	5
A	−248	−115	− 83	−110	−140	+ 30
B	+138	+465	+ 628	+ 20	+213	+133
D	−507	+348	−1360	−270	+ 20	−378

Table 8 — Average Wear Rates for Each Oil

Oil Code	Total Wear, 10^{-7} gm/hr
A	1258
B	2015
D	2508

Table 9 — Effect of Abrasive Particle Size on Filtration

(rate of debris removed to that generated, %)

Filter Type	Abrasive Particle Size, Micron		
	22	15	5
Full-Flow	138	149	137
Bypass	122	86	116
Average	130	118	127

A or D. The explanation for this anomaly is not known, but there was an unusually low wear rate recorded for iron during the abrasive period prior to bypass filtration of Code B oil. However, because the filtration effectiveness was the same for oil Codes A and D, no trend was established and, therefore, oil detergency may not be a factor in preventing filters from removing wear debris from engine oil systems. As in the nonabrasive tests, approximately 70% of the total piston ring was chrome, or face wear.

Since only four of the test runs in Part 2 were repeated, the repeatability of this portion of the program was not established.

Discussion

These experiments confirm the expected benefits to be derived from the use of radioisotope tracer techniques — rapid data collection from a variety of test conditions with a minimum expenditure of effort.

The data collected in the experiments without abrasives in the system are considered to be more amenable to exact interpretation than are the data from the runs with abrasive added. There are several reasons for this. First, the clean system tests were performed during a period of high radioactivity of the piston rings, before the rings had been subjected to the high wear which occurred when abrasives were added to the oil system. This meant that the sensitivity of the measurements was at its greatest during those runs where the rates of wear were lowest. Also, there was no difficulty in cleaning out the oil system between runs to such a degree that no carry-over from run to run was observed. (To overcome this, a new baseline wear rate was established for each run, before abrasive was added.) In the no-abrasive runs fairly constant debris-accumulation (wear) rates were established, and repeatability was good. In the abrasives-added experiments the debris accumulation rates, especially during filtration, did not settle down to a constant rate. The few repeat runs in this phase indicated a lower degree of repeatability.

Although the several factors discussed above were adverse to exact interpretations, the data from the abrasives-added experiments are considered to be useful for relative comparisons of oils, filters, and the like, even though small differences are not to be considered significant.

All of the tests conducted were of very short duration as related to the effective life and dirt-retention capacity of the filters used. Hence, there are no data on the effect of wear-reducing effectiveness of the reduced flow or increased pressure drop in the bypass and full-flow filters, respectively.

The results of this program indicate that detergent-type oils used in a relatively clean engine will effectively minimize piston-ring wear in an engine. Tests of longer duration would be necessary to answer questions concerning wear in relation to filter life and oil detergency level over prolonged periods of time. From the results of these data, it appears that filters are necessary to protect diesel engines from gross quantities of wear debris produced as a result of dirty engine operating environments and to insure that increased engine life will be realized with the use of oils containing detergent additives.

These facts were forceably demonstrated during the dirty engine operating tests. It was shown that, although detergent type oils do permit more wear than straight mineral oils when contaminant materials are suspended in the oil system, filtration (particularly full flow) is effective in reducing oil system wear debris.

For diesel-engine operations that require protection of moving parts against excessive wear caused by abrasive materials, it would be advantageous to use full-flow filters in the lube oil system. When essentially clean conditions prevail, and when using low detergent level oils, bypass filters might be more beneficial in controlling wear rates. However, where engine performance conditions and environments vary through the extremes, best protection would be afforded by the use of both types of filter simultaneously.

Table 10 — Effect of Oil Detergency Level on Filtration

(rate of debris removed to that generated, %)

Oil Code	Filter Type		
	Full-Flow	Bypass	Average
A	178	119	149
B	111	68	90
D	157	127	142
Average	149	105	

Conclusions

The following conclusions are made:

Clean Engine Oil Systems — As the detergency levels of diesel-engine lubricating oils are increased, piston-ring wear rates were reduced, with or without filtration.

Bypass filtration was slightly more effective in reducing piston-ring wear rates than full-flow filtration when using the lower detergent additive level oils. Both types of filtration were equally effective when using the higher detergent level oils.

Filtration was beneficial in prolonging piston-ring life for all of the oil detergent levels studied.

Approximately 70% of the total wear of the piston rings (by weight) occurred in the chromium plating, in sliding contact with the cylinder liner walls.

Abrasive Contaminated Engine Oil Systems — Full-flow filtration was superior to bypass filtration in reducing the contamination and thus engine (piston ring) wear.

On the average, the filters were effective in removing equal amounts of wear debris generated from 5-, 15-, and 22-micron average particle size abrasives.

The 15-micron average particle size abrasive contaminant generally caused greater piston-ring wear than either the smaller or larger average particle sizes.

No trend was established that would indicate that filter effectiveness was reduced with high detergent additive level oils.

The two detergent additive oils apparently dispersed all sizes of abrasive contaminant to the extent that chrome and iron wear of the piston rings with these two oils was approximately twice that with the no-additive oil.

DISCUSSION

Ford Data Compared With Author's Findings

— Robert J. Pocock
Ford Motor Co

WHILE THERE are a few statements in the author's paper which might provoke some disagreement, he is to be congratulated on presenting a very thorough introduction to the measurement of the effectiveness of oil filters as determined by the use of the radioactive tracer technique. It is the first time to my knowledge that this has been done.

I compliment him on his clear definition of this experiment as an evaluation of the effects on piston-ring wear of filtration, oil detergency level, and abrasive particle size. It has been noted that other authors have stated their investigation, using this experimental technique, has shown the effect of some variable on engine wear when the measurement was solely recording the effect on a few piston rings in the engine.

Some laboratory engine tests were made on a Ford, valve-in-head, 6-cyl engine which was operated at 3800 rpm at three-quarter load for 20 hr, with abrasive dust added directly to the crankcase. Fig. A shows the wear of the piston rings by the abrasive particles which were suspended in the oil, passed through a full-flow filter, and were thrown upon the cylinder walls from the crankshaft and the connecting-rod squirt holes. The wear pattern shows the sensitivity of the oil ring to the contaminant in the crankcase oil. Fig. B shows the severity of the wear on the engine bearings and most dramatically evidences the difference in concentration of dust that had been carried through the filter when additive oil had been used as compared to results of a base oil test. In Fig. C, we see a summary of the test results which illustrates the fact that not all areas and components of an engine will react the same to the amount of contaminant carried in the engine oil. This is also true of the effect determined for various contaminant particle size range.

I would like to comment on some of the author's findings and ask him some questions. Why were the three lower compression rings chosen over the two oil rings located near the bottom of the piston skirt? It is known that the oil rings contact many times the volume of oil encountered by the compression rings. Would they not have been a better indicator of the condition of the oil?

It would help if Mr. Halliwell will describe the type of rings used on this engine and whether the extent of ring service and severity of wear could have broken through the chrome clad on the ring face, thereby exposing the cast iron to wear on the ring face as well as the side and back of the ring.

We also believe the value of his work would be increased if the distribution of particle size of each of the specific ratings were known.

The repeatability of the tests of 22.4% is not too impressive. Does the author know if the results would be more reproducible were the quantity of wear material increased over the approximately 16 mg of aluminum oxide added for the 6-hr period? Was the 0.5 gpm of oil that passed through the counter well an adequate sample considering the quantity of oil (10 gpm) being circulated through the engine?

The author points out a trend in his test results indicating that with increasing amounts of detergent the rela-

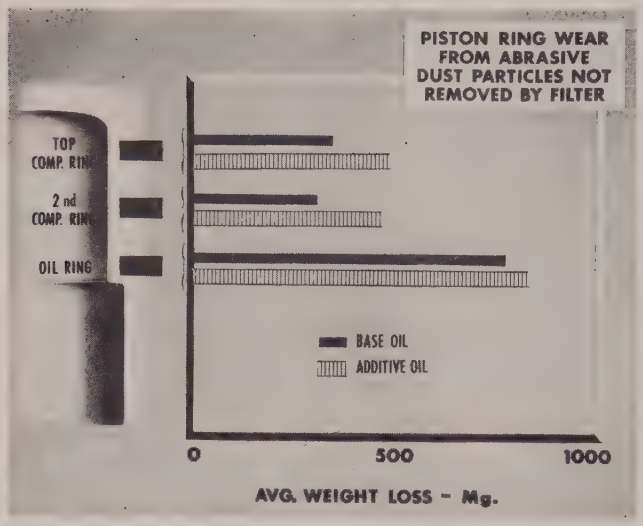


Fig. A

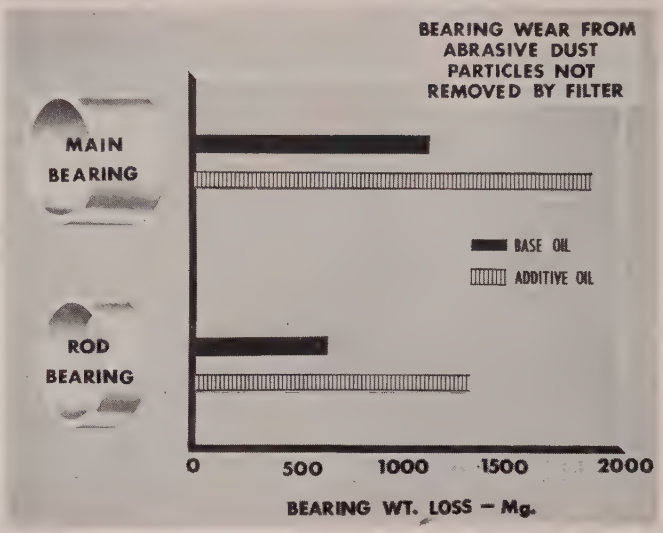


Fig. B

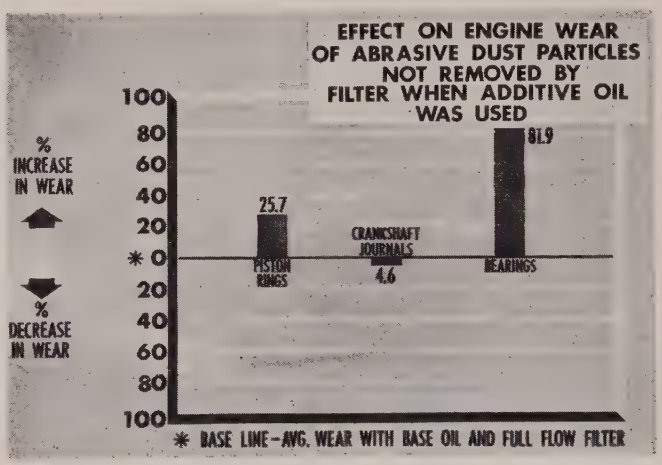


Fig. C

tive effectiveness of the full-flow filter improves. This is contradictory to the Ford test data shown in the illustrations. Mr. Halliwell's trend shown by the part-flow filter is more logical for both types from our experience.

The author refers to the use of the Latin Square design to schedule the runs in Part 1. Is it possible that the wear rate shown for Oil B, for example, was determined at a different time in the service hours of the test piston rings than where the wear rate for each of the filters on Oil B was measured? Why was this system abandoned for the test program in which the abrasive contaminant sizes and the filters were tested with these oils? It would be helpful if the order of the tests was described.

Mr. Halliwell states that "As the detergency levels of diesel engine lubricating oils are increased, piston-ring wear rates were reduced, with or without oil filtration." Let us examine the test results once more, assuming that the test of each oil type in Part 1 was not influenced by any foreign contaminant. The wear of the piston rings occurred during rubbing contact with the walls of the cylinders and the ring grooves. Some wear debris was generated and continued to circulate in the oil. The filters removed some of these particles and, no matter how few were retained, in an absolute condition only a reduction in wear could occur. The reduced wear observed in the tests without filters shows that as the detergent additive level was increased the war rate decreased, demonstrating increased slipperiness or load-carrying capacity of the inhibited oil film. The results were fewer metal-to-metal contacts between ring and bores. If the experience in engine service stopped here, the conclusion that increased oil additive levels decrease piston-ring wear would be true. Unfortunately, engines do not operate under relatively clean conditions in the majority of service applications and, therefore, it must be emphasized that the use of effective filter media to remove dirt as rapidly as possible are ever increasingly important. The trend towards increasing the additive content of oils continues as the more beneficial effects of these compounds are found to be able to offset some inferior quality in the fuel or a part of the engine. We must not be carried away by these localized results. Can we not achieve them with the use of another additive, in place of a detergent, which would not so adversely effect the filter system? The filter is trying to clean the oil of rubbish and abrasives before they are circulated to the areas in which wear will be accelerated by the debris.

In concluding my discussion of Mr. Halliwell's paper, I note that he referred to a procedure for determining if the wear results were being influenced by air-borne dust. Bon-Ami scouring powder was added at a rate of some 3600 times that of the aluminum oxide added, and did not noticeably increase the chrome or iron wear detected.

Radiotracer Technique Presents Problems of Accurate Wear Measurement

— H. R. Jackson
Atlantic Refining Co.

MR. HALLIWELL has presented data which represent a considerable amount of research effort. That this wear test program was no "bed of roses" is indicated by the necessity for a ring change during the investigation because of low activity levels. In addition, as Mr. Halliwell points out, repeatability was particularly bad during the contaminated sump operation. I can, from first-hand knowledge, readily sympathize with him on these problems.

We sincerely question the author's conclusion that his work shows that detergent oils cause more wear under contaminated sump conditions than base oil. We believe it is dangerous to use the radiotracer technique, which depends on constant wear debris suspension characteristics of the lubricant for reliability, when evaluating oils of either suspected or known variable dispersant properties. Here's the paradox. If detergent oils suspend more contaminant (aluminum oxide, in this study) than base oils, which Mr. Halliwell concludes, would they not by inference suspend more radioactive wear debris (iron oxide) than base oils, thereby giving higher indicated wear via the tracer technique, even at the same true wear rate? This problem can only be answered by means of other wear measuring techniques.

We also question some of Mr. Halliwell's statements concerning the filter effectiveness studies. He states that the full-flow filtration was superior to the by-pass filtration in reducing the contamination and attendant engine wear. We ask on what data was this conclusion based? No actual contaminant removal rates were measured, and of course, once the oil filters were cut into the circuit, no meaningful wear rates could be obtained. Only if the radioactive wear debris has the same particle size as the contaminant could the author infer contaminant removal efficiencies. This point is raised because it is difficult to understand how a 40-micron filter can remove 5-micron contaminant so efficiently. One way around this dilemma would be to filter the contaminated oil in a bench filter mock-up, or even on the engine; and then subsequently determine rates of wear over a short period of time on the engine without filtration.

We would like to ask Mr. Halliwell if he can supply information as to the weight per cent distribution of particle sizes in the nominal 5-, 15-, and 22-micron size contaminants? This information would be useful in interpretation of the data.

We would like to bring out one last point, and that concerns the magnitude of wear encountered in these tests. When running on Oil D with no added contaminants, the equilibrium wear rate (chrome plus iron) for the nine rings averages about 200×10^{-7} gm/hr. A reasonable assumption for the "useful life" of a 4¼ in diameter ring is a 2-gm weight less. On this basis, these rings would last approximately 10^6 hr. For those unskilled at mathematics, 10^6 hr is somewhat over 114 years. All this goes to prove that if the Navy could run their diesels as well as Harry Halliwell, he wouldn't have needed to do this work, and I wouldn't have had the opportunity to discuss this interesting paper.

Author's Results Compare Well With Findings at Institute

— George C. Lawrason
Southwest Research Institute

WE HAVE USED the test tool described in this paper for a number of years in the measurement of engine wear. We confirm the soundness of Mr. Halliwell's use of engine and instrument setup and feel certain that his conclusions are well-supported by his test equipment and test method —with one exception until liaison with Mr. Halliwell clears the point. The percentage of total ring weight loss in the GM 3-71 engine that was attributable to other than face wear was much less in the work at the Southwest Research Institute than the 30% that the author found. From the data presented, it is apparent that the weight loss figures were not compiled from tests that included abrasive introduction, and, therefore, no unusual conditions were present. It would appear that fuel sulfur content may be responsible for an increased ratio of face-to-side wear, and most of the work at the Institute was done with 1% sulfur content fuel. Since the cylinder liner is largely affected by increased fuel sulfur content, it would appear that the face wear-side wear ratio might be higher. At any rate, several careful spectro-analyses of the drain oils indicated no measurable ring wear debris other than chrome from the face. This fact was confirmed by measuring the hot rings.

Further, the author's paper appears to be concerned only with abrasive wear when, in fact, corrosive wear is always present. Is it conceivable that the introduction of abrasive material accelerated corrosive ring wear also?

The author maintains that 15-micron particles were generally more damaging than the 5- or 22-micron particles. It would be interesting to speculate why, and this speculation gives rise to two fairly obvious facts. First of all, the same quantity of each size abrasize was introduced every time. In the case of the 22-micron particles, a larger percentage would tend to be centrifuged out of the oil system or hang up in fitting joints so that they cannot reach the piston rings; however, relatively few 22-micron particles need reach the ring belt to cause severe damage. Conversely, in the case of the 5-micron particles, distribution should be relatively better in the oil system and more of these particles would tend to reach the ring belt. Perhaps this may account for the reason that wear rates do not vary directly with the abrasive particle size.

This is the type of paper that stimulates creative thinking which in turn may lead to additional effort in the field of engine wear. We feel that Mr. Halliwell is to be congratulated on contributing an interesting paper to the field of engine wear.

Other comments on this paper will be found on p. 303, following the Robbins-Pinotti-Jones paper.

Author's Closure To Discussion

IN ANSWER to the questions by Mr. Pocock, the Bon-Ami was introduced into the intake air ahead of the engine air cleaner. The intended purpose was to give some indication as to whether normal dust conditions in the test cell might have contributed to the ring wear over the extended period of test. Thus, this test phase was essentially an evaluation of the air cleaner's effectiveness.

The piston rings used were standard General Motors parts for the Model 3-71 diesel engine. They were steel rings, chromium-plated and taper-faced. Visual inspection of the rings at the end of test gave no evidence that the chromium had worn off completely in any area of the ring surface.

The particle size distribution of the abrasives used in this program was determined by counting the particle distribution in 20 microscopic fields and averaging the counts obtained. The results are shown in Table A.

A low abrasive contaminant rate was used during the tests so that engine wear would not be unduly high and the test could be thereby extended to survey additional effects. During the initial tests, higher contaminant addition rates were used, but results were inconclusive and there was no definite evidence that test reproducibility was improved thereby.

The lubricating oil sample size is considered to be entirely adequate. The flow rate was sufficient to produce three changes per minute of the 560-ml capacity of the counter well. No significant difference was observed when oil sample flow rates as high as 2½ gpm were used.

The piston oil control rings might have been a better indicator of wear associated with lubricating oil conditions than the three lower compression rings. However, satisfactory compression ring operation was considered the more critical factor and the tests were, therefore, made on that basis.

The paper states that a randomized block of the Latin Square design was used to establish the order of test runs during Part 1 of the investigation. A more accurate terminology for the arrangement used would be a randomized block of the factorial design with three replicate blocks being used during this phase. An example of the test order arrangement is given in Table B. The numbers refer to the order of test.

The test runs during Part 2 of the investigation were randomized with regard to test oils in an attempt to minimize effects due to the time factor. The factorial plan was abandoned because it was considered that the life of other engine components — that is, bearings, cams, and the like — would be reduced by the addition of abrasives to a level where catastrophic failure might occur.

With reference to Mr. Jackson's comments, it is agreed that, for reliable results, the radiotracer technique depends on the constant suspension of the wear debris in the lubricating oil. However, during our base oil tests, rigid monitoring schedules did not indicate any significant sedimentation.

It is apparent, from the slopes of the curves for the 12–18 hr test period in Figs. 2–7, that the full-flow filters were, on the average, more effective than bypass filters in removing radioactive piston-ring wear debris. These data were intended only to show the relative effectiveness of the two filter types in removing radioactive piston-ring wear debris produced as a result of the abrasive action of the three size grades of abrasive; therefore, no attempt was made to measure actual contaminant removal rates. Since the abrasive contaminant was not radioactive, we could not discriminate by radioactive means between the particle sizes trapped by the filter.

As noted by Mr. Jackson, the rings would have a very long life at the reported wear rate. The wear rate was intentionally kept low by low contaminant addition rates so that a relatively large number of radioactive tests could be made before the total wear reached a point requiring ring replacement. Another factor contributing to the low wear rates was the use of a new lubricating oil charge for each test during the clean sump operations. Thus, during the entire test period the oil additive agents would be functioning at their maximum effectiveness, and there would be none of the accelerated wear rates normally occurring near the end of the oil's useful life.

Concerning Mr. Lawrason's comment regarding the proportions of iron and chromium wear, it is probable that the Southwest Research Institute tests, in which the top ring only was radioactivated, reflected fuel effects to some extent. The Experiment Station results, in which the three lower compression rings were activated, probably were more related to lubricant effects. This might explain the relative differences in the proportion of iron and chromium wear reported in the two investigations.

Table A — Actual Abrasive Particle Size Distribution, %

Size Range	Nominal Particle Size, Microns		
	22	15	5
Larger than 25 microns	7.8	2.4	< 1
20–25 microns	10.6	5.0	2.7
15–19 microns	18.9	11.4	2.4
10–14 microns	20.7	21.6	11.1
5–9 microns	24.0	26.8	19.6
Smaller than 5 microns	18.0	31.9	64.0

Table B — Test Order Arrangement

Lubricating Oil	Filter System Used		
	None	Full-Flow	Bypass
Code A	6	12	5
Code B	3	9	7
Code C	8	11	4
Code D	2	10	1

the use of Radioactive Tracer Techniques to determine the effect of operating variables on Engine Wear

B. A. Robbins, General Metals Corp.
P. L. Pinotti, Standard Oil Co. of California
D. R. Jones, California Research Corp.

THE ENTERPRISE Division of General Metals Corp., Standard Oil Co. of California, and California Research Corp. joined in a cooperative research program to determine the effects of operating variables on wear in a medium-speed diesel engine. The radioactive tracer technique[1] was chosen for this program because it is ideally suited for any investigation of wear in which a large number of variables are to be studied in a relatively short period of time. In this study the effects of engine speed, load, jacket water temperature, fuel temperature, and chromium-plated piston rings were investigated using a distillate and two residual fuels. The results obtained in these tests are discussed in the following sections.

Test Engine

The engine used in this program was a 3-cyl Enterprise Model DSM-3 of 8-in. bore and 10-in. stroke and rated at 162 hp at 1000 rpm. It is a 4-stroke cycle, trunk-type engine with an open combustion chamber. For all tests the engine was equipped with cast-iron cylinder liners. Each piston had three compression and two oil control rings. A radioactive ring was installed in the top compression ring position of each piston. Fuel injection nozzles used in all tests were Bosch Model ADL 130T30 having six 0.010 in. diameter holes with a 130-deg spray angle.

The engine was loaded by means of a cradled generator coupled directly to the engine. Conventional instrumentation was used to measure engine load, speed, temperatures, and fuel flow rates. A photograph of the test installation is shown in Fig. 1.

Radioactive Tracer Techniques

Cast-Iron Rings — In the first part of the program, irradiated cast-iron compression rings were used. For this work a conventional Geiger tube counting system proved entirely satisfactory. The system is shown schematically in Fig. 2. It consisted of an electrically driven pump which continuously circulated oil from the crankcase through a counting chamber containing the Geiger tube and then back to the crankcase. The Geiger tube was

[1] "Application of Radioactive Tracers to Improvement of Automotive Fuels, Lubricants, and Engines," by P. L. Pinotti, D. E. Hull, and E. J. McLaughlin. SAE *Quarterly* Transactions, Vol. 3, October, 1949, pp. 634–638.

connected to a ratemeter which calculated the rate at which impulses were received from the Geiger tube. The output of the ratemeter was fed to a chart recorder which plotted count rate against time. Since count rate is directly proportional to the concentration of radioactive iron particles in the lubricating oil, the slope of the count rate versus time plot on the chart recorder was directly proportional to the rate at which iron was worn from the irradiated rings.

To obtain absolute wear rates in mg/hr, it was necessary to measure the count rate of a standard solution containing a known amount of dissolved iron previously removed from the irradiated rings. This measurement was made in a counting chamber of the same size and shape as used on the test engine. The measured activity of the standard solution was converted to specific activity, counts/sec/mg of iron. Knowing the specific activity, absolute wear rates were calculated using the following equation:

$$\text{Wear rate (mg Fe/hr)} = \frac{\text{Chart recorder slope (cts/sec/hr)}}{\text{Specific activity (cts/sec/mg Fe)}} \times \frac{\text{Volume of oil in crankcase}}{\text{Volume of counting chamber}}$$

The specific activity of the irradiated rings decreased with time in accordance with the decay rate of the radioisotope, Fe-59. The half-life of Fe-59 is 46 days. The half-life of the rings used in these tests proved to be slightly longer (47 days), probably due to the presence of minute amounts of cobalt-60.

Chromium-Plated Rings — In the second phase of the program, irradiated chromium-plated compression rings (face only plated) were installed in the top ring groove of each piston. In these tests, samples of lubricating oil were drawn from the crankcase at scheduled intervals during each run. Each sample was placed in a special chamber and counted with a scintillation crystal. This instrument is several times as sensitive as a Geiger tube and has the added advantage of emitting electrical impulses having voltages which vary with the energies of the gamma rays it receives. By using a suitable voltage discriminating device to count impulses from the scintillation crystal, it is possible to determine the relative amounts of radioactive iron and chromium present in an oil sample.

For these tests, the output from the scintillation crystal was fed to a gamma ray spectrometer and the counting operation was carried out in two steps. First, the spectrometer was adjusted to count over a narrow voltage range around 0.32 mev which is the principal gamma energy peak for chromium-51. After obtaining an accurate count rate at this condition, the spectrometer was set to count all impulses above about 0.4 mev. With this procedure the maximum number of counts were obtained on iron-59 including its two principal gamma energies of 1.1 and 1.3 mev. Chromium-51 emits no gamma rays having energies above 0.4 mev. To convert count rates measured at these two conditions into actual amounts of chromium and iron present in each oil sample, it was necessary to count standard

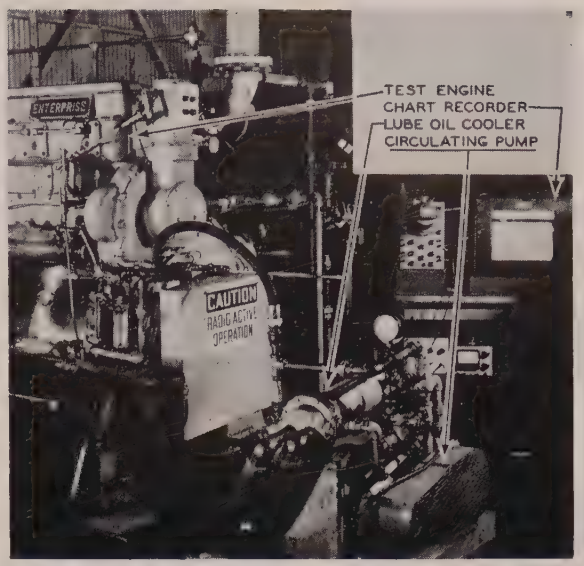


Fig. 1 — Test installation

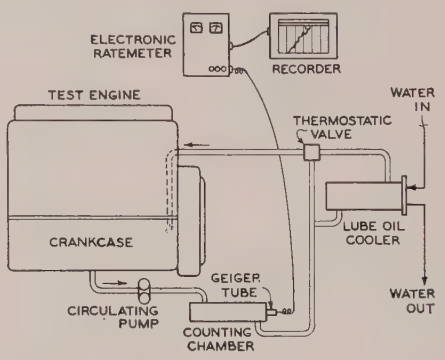


Fig. 2 — Geiger tube counting system used on test engine

RADIOTRACERS were used to study the wear effects of engine speed, load, jacket water temperature, fuel temperature, and chromium-plated rings in a medium-speed diesel engine. One distillate fuel and two residual fuels were tested. This paper describes the tests and their results. Some of the conclusions are:

- **The brake thermal efficiency with high viscosity residual fuel was essentially equal to distillate diesel fuel over a wide range of loads, providing the residual fuel was heated to the proper temperature.**
- **Engine speed did not affect the wear rate of cast-iron rings when distillate fuel was used, while with residual fuel wear decreased with increased speed.**
- **With distillate fuel, engine load had essentially no effect on cast-iron ring wear. With residual fuel, decreasing engine load produced a marked increase in ring wear.***

* Paper presented at SAE Summer Meeting, Atlantic City, June 16, 1959.

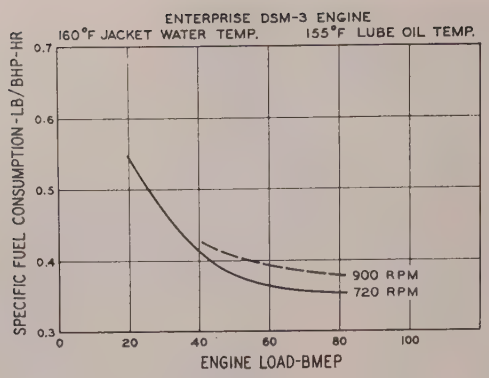


Fig. 3 — Specific fuel consumption — Fuel A

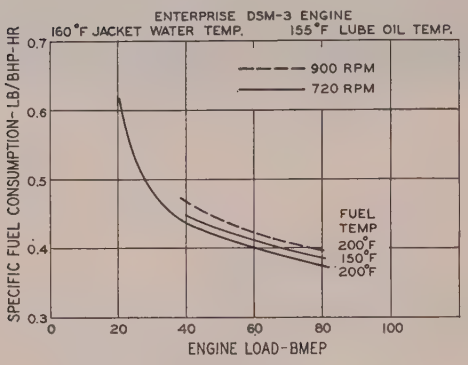


Fig. 4 — Specific fuel consumption — Fuel B

solutions of chromium and iron in the same container and under the same conditions. Since iron-59 also emits gamma rays below 0.4 mev, a correction had to be applied to the chromium-51 counts to adjust for the iron-59 radiation in this lower energy region.

Each sample of oil was counted by this method. Knowing the sample size in relation to the quantity of oil in the crankcase, it was a simple matter to calculate the individual wear rates of chromium from the face of the rings and iron from the sides of the rings.

Ring Weight Loss Material Balance

The irradiated cast-iron rings were weighed before and after the wear tests conducted with these rings. The combined weight loss of the three rings was compared with the weight loss calculated from the radioactive wear data. The results obtained were as follows:

Total gravimetric weight loss, gm	13.57
Calculated weight loss from radioactive measurements, gm	11.74

It is of interest that more than 85% of the total wear on these rings could be accounted for by radioactive measurements.

Wear Test Procedure

The engine was operated for 6 hr at each test condition. This proved to be sufficient time to obtain accurate wear rates. To avoid possible carry-over effects from one run to the next and to keep the radioactivity in the lubricating oil as low as possible, the oil was changed at the end of each 6-hr test. The engine was run without lubricating oil filters to ensure that all of the radioactive wear debris remained in the lubricating oil.

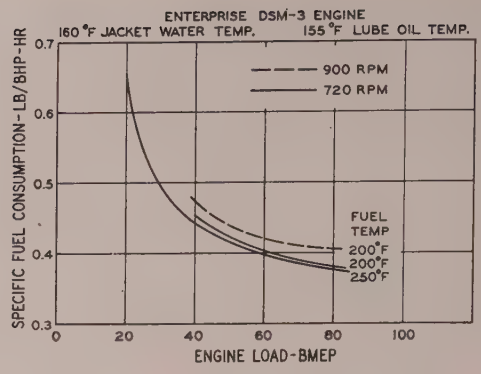


Fig. 5 — Specific fuel consumption — Fuel C

Most of the data were obtained with the engine operating at 720 rpm. A few runs were made at 900 rpm with both distillate and residual fuels to determine the effect of speed on wear. In each series of tests, data were obtained at loads of 20, 40, 60, and 80 bmep. For all tests, inlet lubricating oil temperature was held constant at 155 F, and inlet jacket water temperature was maintained at 160 F except when this variable was being studied.

Test Fuels and Lubricating Oil

Three fuels were selected for these tests. Fuel A was a distillate diesel fuel meeting ASTM 2-D specifications. Fuel B was a residual fuel of intermediate viscosity (573 SSU at 100 F) which met ASTM No. 5 fuel oil specifications, and Fuel C was a high viscosity (6044 SSU at 100 F) residual fuel meeting ASTM No. 6 fuel oil specifications. Except for sulfur content these three fuels covered a wide range of properties. Although the three fuels were commercially available products, they were carefully selected to have comparable sulfur contents so that sulfur would not be a factor in evaluating the wear characteristics of the three fuels. Complete inspections on these fuels are contained in Table 1.

The lubricating oil used in all tests was a commercially available marine diesel-engine lubricating oil of SAE 30 grade. Inspections of this oil are shown in Table 2.

Preliminary Engine Tests

Before starting the radioactive ring tests, it was considered advisable to obtain performance data with the three fuels to determine if there were any significant differences in their combustion characteristics. Figs. 3–5 show specific fuel consumptions over a range of loads for Fuels A, B, and C, respectively. It will be seen from Figs. 4 and 5 that fuel temperature had an effect on performance of the two residual fuels. The lower bsfc's obtained at higher fuel temperatures probably reflected better atomization due to reduced viscosity. Fig. 6 is a plot of specific fuel consumption in terms of Btu per bhp-hr against engine load and shows that on a Btu input basis all three fuels performed essentially alike. Of particular interest was the finding that even the No. 6 fuel oil (Fuel C) burned well

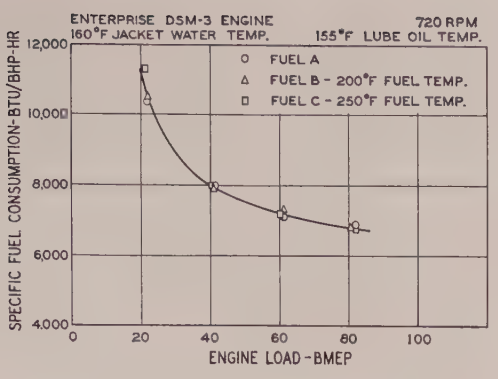


Fig. 6 — Specific fuel consumption of three test fuels based on Btu input

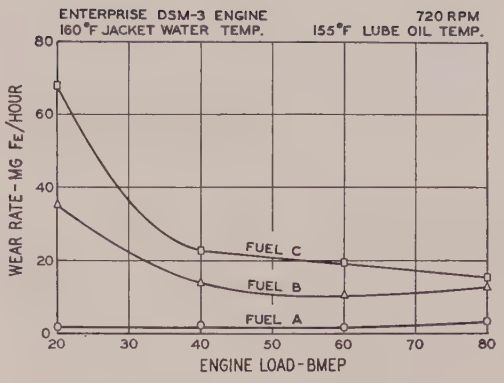


Fig. 7 — Effect of load on piston-ring wear

over a wide range of loads and gave essentially the same brake thermal efficiency as did the distillate fuel (Fuel A). With this background, we felt confident that any differences in wear characteristics of the three fuels could not be attributed to gross differences in their combustion characteristics.

Wear Tests with Irradiated Cast-Iron Compression Rings

In this portion of the program wear rates of the three fuels were compared at a number of different operating conditions using standard cast-iron compression rings. For all tests on Fuel B, injection pump inlet temperature was held at 200 F. Except for tests on the effect of fuel temperature, Fuel C was tested at 250 F. In the first series of tests the engine was operated at 720 and 900 rpm to determine if distillate and residual fuels reacted differently to engine speed. Table 3 shows that with Fuel A, speed had no effect on wear at high load. In contrast, the wear rate of Fuel C was appreciably lower at 900 rpm than 720 rpm. The lower wear observed at 900 rpm may have been due to better combustion resulting from higher turbulence, higher mean cycle temperature, and improved injection characteristics.

In studying the effect of engine load on wear, we again found a marked difference between the distillate and residual fuels. From Fig. 7 it will be seen that load had an almost negligible effect on the wear rate of Fuel A, whereas the wear rates of Fuels B and C increased markedly as load was reduced. Although the earlier performance tests showed no differences in the gross combustion characteristics of the three fuels, these data suggest that minor amounts of the residual fuels did not burn completely at lighter loads resulting in higher wear with these fuels.

It is well-known that jacket water temperature has a significant effect on wear, particularly with high sulfur fuels. In these tests it was found that over a range of jacket water temperatures from 140–200 F the wear rate with Fuel A did not change significantly. It is likely that some minimum jacket water temperature exists for Fuel A in this engine below which a marked increase in wear rate would occur. In tests with Fuel B over the same jacket water temperature range, it was found that an increase in wear occurred below 160 F. These results are shown on Fig. 8. One reason for the different behavior of Fuels A and B to changes in jacket water temperature may be that the reduced temperatures adversely affected combustion of Fuel B near the

Table 1 — Test Fuel Inspections

	Fuel A	Fuel B	Fuel C
Fuel Type	Distillate	Residual	Residual
ASTM Specification No.	2-D	5	6
Gravity, deg API	32.0	11.9	8.3
Specific Gravity	0.8654	0.9868	1.0122
Viscosity			
SSU at 100 F	39.1	573	6,044
SSU at 210 F	1.45c/s	58.6	141
SSF at 122 F	—	28.0	213
Sulfur, weight %	0.83	0.86	0.94
Conradson Carbon, weight %	0.05[a]	14.6	15.6
Ash, weight %	0	0.05	0.03
Sediment and Water, volume %	Trace	0.2	0.2
Sediment by extraction, weight %	0.03	0.09	0.09
Cetane No.	43.5	—	—
Carbon, weight %	85.76	86.91	87.15
Hydrogen, weight %	13.08	10.34	9.95
Gross Heating Value, Btu/lb	19,495	18,243	18,090
Gross Heating Value, Btu/gal	140,480	149,920	152,500
Net Heating Value, Btu/lb	18,200	17,273	17,160
D 158 Distillation, F			
Start	392	348	442
10% recovered	472	462	560
50% recovered	550	[b]	[c]
90% recovered	620		
End point	673		

[a] Measured on 10% bottoms.
[b] Cracking at 648 F, 47% recovered.
[c] Cracking at 674 F, 34% recovered.

Table 2 — Test Lubricating Oil Inspections

SAE Grade	30
Viscosity	
SSU at 100 F	519
SSU at 210 F	58.4
Viscosity Index	62
Flash Point, F	430
Ash, weight %	0.41
Conradson Carbon, weight %	0.67

Table 3 — Effect of Engine Speed on Wear Rate

Engine Speed, rpm	Wear Rate, mg Fe/hr	
	Fuel A	Fuel C
720	3.2	15.6
900	3.2	10.0

cylinder walls to a considerably greater extent than Fuel A.

In burning residual fuels, it is necessary to supply fuel at the proper viscosity to the injection nozzles. This is accomplished by heating the fuel to reduce its viscosity. The effect of fuel temperature, as measured at the injection pump inlet, was studied with Fuel C to determine whether this operating variable had a significant effect on wear. It will be seen from Fig. 9 that at fuel temperatures below 225 F wear rate increased markedly. This temperature corresponds to a fuel viscosity of about 100 SSU at 100 F. As temperature was increased above 225 F, no significant further reduction in wear was observed indicating that fuel viscosities less than 100 SSU at the pump were not required with the particular injection system used on this engine.

Wear Tests with Irradiated Chromium-Plated Compression Rings

In the second phase of the program tests were conducted with irradiated chromium-plated rings (face only) installed in the top ring groove of each piston. Standard nonradioactive cast-iron rings were used in the remaining compression ring positions. As mentioned above, a gamma ray spectrometer was used in conjunction with a scintillation crystal to measure the amounts of metal worn from the ring face and the ring sides simultaneously.

Comparison of the data obtained with chromium-plated rings and cast-iron rings is shown in Fig. 10. It will be noted that a large reduction in wear was obtained on all three fuels with the chromium-plated rings. It is also of interest that with all three fuels about half of the material worn from the rings was chromium and came from the ring face and about half was iron and came from the sides of the ring. This shows that ring side wear is an appreciable part of total ring wear and that with residual fuels, which cause a large increase in wear, some means should be considered for reducing side ring wear as well as face ring wear.

Another difference in the wear characteristics of chromium-plated rings compared to cast-iron rings was the effect of engine load on wear with residual fuels. This is shown in Fig. 11. With Fuel A, wear rate of the chromium-plated rings was not affected by load as was the case with cast-iron rings. With Fuel B, however, wear rate of the chromium-plated rings did not increase as load was reduced which suggests that chromium has a remarkable ability to resist the wear-producing materials formed when residual fuels are burned at light loads in this engine.

Summary

The results obtained in this radioactive ring wear test program may be summarized as follows:

1. The brake thermal efficiency with high viscosity residual fuel was essentially equal to distillate diesel fuel over a wide range of loads, providing the residual fuel was heated to the proper temperature.

2. Engine speed did not affect the wear rate of cast-iron rings when distillate fuel was used, while with residual fuel wear decreased with increased speed.

3. With distillate fuel, engine load had essentially no effect on cast-iron ring wear. With residual fuel,

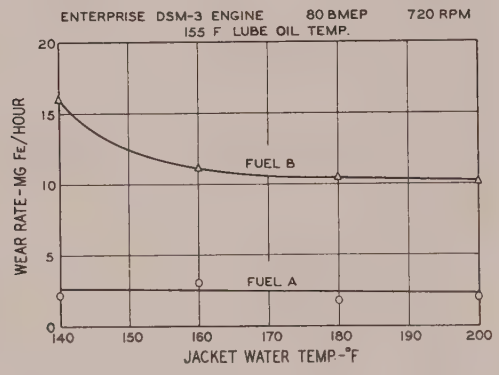


Fig. 8 — Effect of jacket water temperature on piston - ring wear

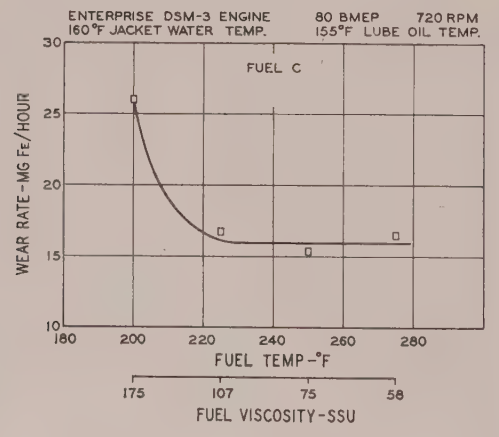

ENTERPRISE DSM-3 ENGINE 80 BMEP 720 RPM
160°F JACKET WATER TEMP. 155°F LUBE OIL TEMP.
WEAR RATE - MG Fe/HOUR
FUEL C
FUEL TEMP.-°F
FUEL VISCOSITY-SSU

Fig. 9 — Effect of residual fuel temperature on piston - ring wear

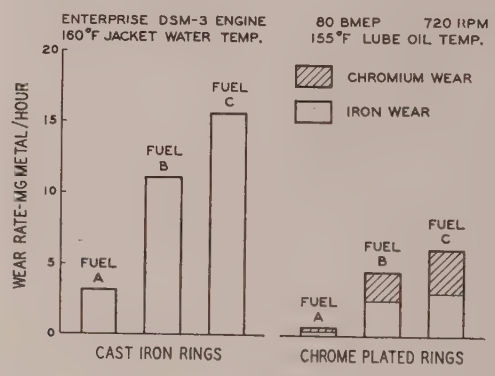


Fig. 10 — Comparison of wear rates with cast-iron and chromium - plated piston rings

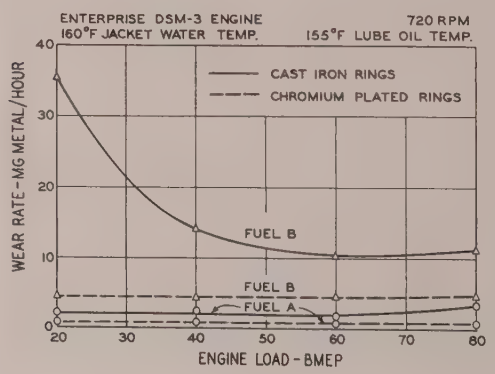

ENTERPRISE DSM-3 ENGINE 720 RPM
160°F JACKET WATER TEMP. 155°F LUBE OIL TEMP.
CAST IRON RINGS
CHROMIUM PLATED RINGS
WEAR RATE - MG METAL/HOUR
FUEL B
FUEL A
ENGINE LOAD - BMEP

Fig. 11 — Effect of load on piston - ring wear with cast-iron and chromium - plated rings

decreasing engine load produced a marked increase in ring wear.

4. Cast-iron ring wear rate with distillate fuel did not change over a jacket water temperature range of 140–200 F. With residual fuel, wear increased at temperatures below 160 F.

5. With residual fuel, cast-iron ring wear increased below a certain minimum fuel temperature which will vary depending upon fuel viscosity.

6. Chromium plating the face of piston rings markedly reduced the wear with both distillate and residual fuels.

7. With residual fuel, wear rate did not increase as load was reduced when chromium-plated rings were used.

8. With both distillate and residual fuels the amount of iron worn from the sides of the rings was about equal to the amount of chromium worn from the face of the rings.

Acknowledgment

A number of people in each of the companies participating in this program contributed a great deal of time to this project. The authors, however, particularly wish to acknowledge the contribution of C. D. McLean of California Research Corp. to the overall success of the program.

DISCUSSION

Questions Effects of Jacket Temperatures Above 200 F

— T. O. Kuivinen

The Cooper-Bessemer Corp.

THE AUTHORS have presented a very interesting study of determination of wear rates by using radioactive piston rings. This work demonstrates a procedure through which considerable data can be obtained in a shorter period of time than usually required in wear research. The paper very effectively reports the results and attempts to explain some of the behaviors noted.

Fig. 8 showed the results of varying jacket water temperature. We are disappointed that the authors chose to stop this test at a maximum of 200 F water temperature. Installations are made today with pressure on the jacket so that water temperatures as high as 250 F are often encountered. Such installations are usually made for the purpose of supplying waste heat for other purposes within the plant. It would have added much to the knowledge if we knew what would happen to the wear rates as temperatures were increased above 200 F. Do the authors have any supplementary data to add in this area?

I was most interested in the observation that the wear rate at 900 rpm using residual Fuel C was less than that found at 720 rpm. This is contrary to what I would guess would occur and the authors attempted to explain this by suggesting higher turbulence, higher mean cycle temperature, and improved injection were the reasons. Could they give further explanation as to how these three factors were improved by the increase in speed?

I have one further question. As jacket water temperature was increased to observe the effect on engine wear it is noted that lube oil temperature was maintained constant. Did the heat load to the lube oil increase with increased jacket water temperature? I presume that the same lube oil was used throughout this jacket water temperature range. Is this a correct assumption?

We hope that this work will encourage others to do wear research in this manner, for the reporting of such research adds greatly to the experience knowledge on the effects of using the economy residual fuels.

Wear Particle Loss — Factor in Engine Wear Tests

— C. K. Murphy and R. McClintock

General Motors Corp.

(This discussion covers Mr. Halliwell's paper found on p. 288, as well as the Robbins-Pinotti-Jones paper.)

WE WERE PARTICULARLY interested in a statement by Messrs. Robbins, Pinotti, and Jones that "more than 85% of the total wear . . . could be accounted for by radioactive measurements." This statement implies that almost 15% of the total wear could not be accounted for by radioactive measurements. This latter 15% is of special interest to us in that we at General Motors Research have also encountered this phenomenon which we term "wear particle loss" while using radioactive tracer techniques in studying hydraulic valve lifter wear in a 1958 V-8 engine.

The detection system used was similar to that described by Mr. Halliwell in his paper. It was a continuous-type counting system in which the engine oil was pumped continuously through a 1-qt aluminum counter well. The counter well contained a scintillation detector, which was a thallium activated sodium iodide crystal. The count rate was plotted continuously while the engine was running and was converted into wear values by counting a standard solution in an identical counter well.

When the radioactive wear values were compared to actual wear data obtained by weighing the irradiated lifters before and after the tests, some discrepancies were noted. In each test the radioactive wear figures were lower than the gravimetric wear determinations. This "wear particle loss" was not constant, however, but varied from 14% to 43% as can be seen in Table A.

The type of test in the table refers to the engine running condition while the observed wear is the total wear obtained with radioactive tracer techniques. There are three continuous tests and three cyclic tests reported.

In general, the continuous tests were representative of mild valve lifter wear conditions, while the cyclic tests were conducted with lubricant and test conditions such that "severe" or catastrophic hydraulic valve lifter wear would occur. Tests 5 and 6 were terminated after 16 and 8 hr, respectively, due to valve lifter failure.

Not only was there as much as a 43% "wear particle loss" obtained but in one test a "negative wear rate" was also observed. While running the engine at a constant speed and continuously recording the radioactivity in the oil, it was noted that the wear particles in the oil began

Table A — Wear Particle Loss

Test No.	Test Type	Test Hours	Gravimetric Wear, mg	Radioactive Wear, mg	Wear Loss, %
2	Continuous	51	12.5	7.6	39
3	Continuous	150	43.6	25.2	42
4	Continuous	22	57.6	33.0	43
	Cyclic	5	20.8	15.9	24
5	Cyclic	16	635.6	506.4	20
6	Cyclic	8	408.0	350.0	14

to decrease as indicated by a decrease in the count rate. In Fig. A the total wear is plotted against the running time. Since the slope of this curve is the wear rate, a decrease in the wear particles in the oil indicates a "negative wear rate." A total wear of 32 mg was obtained after 1 hr of operation. Subsequently, the total observed wear in the oil decreased for the next 6 hr operation. In addition, it required a total of 20 hr to obtain the same total observed wear as was obtained at the end of the first hour. The decrease in activity between 1 and 6 hr indicates "wear particle loss" which occurred in a continuous type test.

During the cyclic tests wear particle loss not only occurred in a running engine but was also obtained while the engine was shut off. Each cycle consisted of running the engine at 2500 rpm for 10 min after which the engine was shut off for 50 min. The oil was circulated continually through the counting and heat exchanger system. A baseline count rate was obtained before each engine start with a continuous count rate being recorded during the complete cycle.

Fig. B presents the wear curves for three of the 16 cycles of Test 5. In this slide the count rate is plotted against time. Cycles 3 and 7 show the type of wear curves which might be expected in that the maximum count rate was observed at the end of the 10-min run with no indication of "wear particle loss" during the run. However, during the shutoff time the count rate decreased, thus indicating a "wear particle loss" even when the engine is not running. This phenomenon was evident in 13 of the 16 cycles. In four of the cycles the "wear particle loss" appeared as negative wear rate during the 10-min run and continued at the same apparent rate until the engine was restarted. Cycle 11 is typical of this and shows that maximum count rate was obtained after 3 min of running.

We would like to hear the experiences of others who have used radiotracer techniques to study engine wear. We would like to know if anyone else has experienced this lack of correlation between radioactive and gravimetric wear, or if they have obtained negative wear. If so, how have they compensated for or eliminated the problem.

Even if one were to talk only about relative wear rates, wear data could be seriously in error if the wear particle loss is not completely understood.

I would like to ask the authors to explain their wear particle loss, if possible, and to indicate what compensations were made for this phenomenon, if any. I would also like to ask Mr. Halliwell if he has encountered wear partice loss or negative wear.

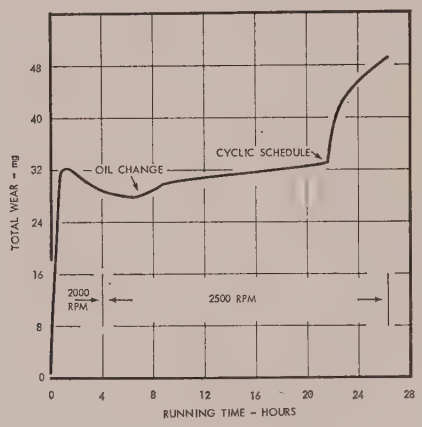


Fig. A — Hydraulic valve lifter wear lubrited foot

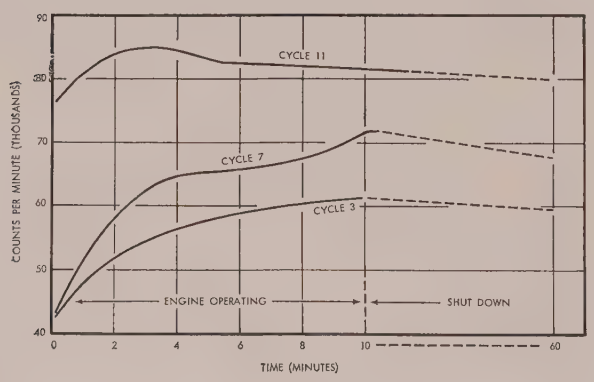


Fig. B — Cyclic wear

Robbins-Pinotti-Jones Closure to Discussion

IN ANSWER to the questions by Mr. Kuivinen, we were prevented by the test work load from studying the effects of jacket water temperatures higher than 200 F. This might be a good field for further study. However, our wear curve appears to become asymptotic at about 190 F and higher temperatures may not, therefore, cause a decrease in wear rates.

The lower wear rate at 900 rpm, as compared with that at 720 rpm, was unexpected. We have no explanation for it beyond that offered in the paper. We do not know whether the lubricating oil heat load increased. The same lubricating oil was used throughout the test.

Regarding Mr. Murphy's inquiry, no information is available that would account for the 15% of the total wear debris not indicated in the lubricating oil. Possibly some or all of the unaccounted debris passed out with the exhaust gases, but we have no measurements to substantiate this. In some tests, there was a rather wide scatter of data, but no negative wear was observed in any instance.

ORAL DISCUSSION

of Halliwell and Robbins-Pinotti-Jones Papers

Reported by Alan R. Schrader

U. S. Naval Engineering Experiment Station

Walter C. Arnold, Fairbanks, Morse & Co.: If the primary interest is cylinder linear wear, isn't the ring wear data obtained of limited usefulness? Is there any information available as to the relative wear rates of piston rings and cylinder liners? Can an estimate be given as to the relative sensitivity of Geiger and scintillation counters?

Radioactive tracer tests were made by Fairbanks, Morse & Co. to determine the proportions of the wear debris passing into the exhaust gas and lubricating oil of a 2-stroke cycle diesel engine. From 10–40% of the wear debris was carried away in the exhaust gases, with the actual proportion depending on the engine load and other factors. The total wear determined by radioactive tracer methods agreed very closely with the wear as determined by physical measurements.

Mr. Robbins: Ring and liner wear have been found to have close correlation. This has been substantiated by field experience.

Robert J. Pocock, Ford Motor Co.: Fleet tests made by Ford have shown very excellent correlation between the wear rates of cylinders and piston rings. Any factor that affected ring wear had a similar effect, in about the same proportion, on cylinder wear.

Harry Halliwell, U. S. Naval Engineering Experiment Station: Negative wear rates were observed during some of the early radioactive tracer studies at the Experiment Station. The cause was found to be the sensitivity of the scintillation crystal to temperatures over 135 F. When the lubricating oil in the counter well was maintained at a lower temperature, no further difficulty of this nature was experienced.

We have not made measurements of the relative sensitivity of the Geiger and scintillation counters, but I have seen literature references indicating that the scintillation counter is of the order of 20 times more sensitive than the Geiger unit.

Computer-Based Selection of Balanced-Life Automotive Gears

G. J. Huebner
Chrysler Corp.

TODAY, GEAR DESIGN, and engineering in general, is undergoing a revolution. The computer has come upon the scene and, instead of replacing people, as we might have expected, is being used to accelerate scientific progress and to learn more and more about the present fields of engineering. It allows us to tackle problems for which we had never previously had time, and to evolve optimum designs. I will present part of a series of gear design charts that are the result of a computer program which could occupy an engineer from now until about February, 2012 A.D. if he were to calculate them without computer aid.

A good example of the increase in speed possible through the use of a computer is seen in our hypoid gear calculations. Any gear designer recognizes that these are time consuming. Working under pressure for two weeks, a man can calculate three of these designs for production, but one can hardly hope to keep this up except under conditions of extreme urgency. The computer, on the other hand, solves one of these problems in 14 min. This is not just an extreme example of how fast the computer can do a job, for there are many other instances in which the ratio of manual-to-electronic calculating speed is even more heavily weighted in favor of the computer. These are examples of bread-and-butter

THIS PAPER describes the balanced-life concept of gear design — in which the gear and pinion are designed to fail simultaneously. An example is presented to show how this concept allows a combination of minimum size and maximum capacity in gas turbine application. Various reasons for failure and factors in long gear life are discussed. The author analyzes the calculations needed and their programming for a digital computer.

Calculating gear designs for production is a time consuming, demanding task to do manually. The use of the computer has changed this — and brought about better gear design by making it possible to study more detailed analyses to evolve optimum solutions.

This paper was the Sixth L. Ray Buckendale Lecture, presented at the 1960 SAE Annual Meeting.

When I was in school, taking an occasional course involving gear design, it was my opinion that gear problems were solved by descriptive geometry and a few canned equations, which involved an endless series of empirical assumptions. I didn't think I would like to be a gear designer and be forced to work in such an unscientific field. In the years that followed I learned that I had touched the field of gear design only enough to be able to use the right words when I spoke to gear designers — specialists in an area of mechanical engineering that I knew little about.

As I became more familiar with gears, I began to recognize them not as cog wheels, but as cams — uniform velocity cams, at that — arranged so that the load is transferred to a succeeding pair before the first pair is out of contract. I have seen a number of other drive methods suggested, and I have been involved in the design of some which avoided gears entirely; but I have never found any mechanism that is, as compact, efficient, and versatile as a gear.

Gears are light in weight and dependable. They adapt themselves well to changes in environment, operating at a wide range of temperatures, under conditions of alternate flooding and starving of lubricant, and yet they continue to operate for millions of cycles without failure, slip, or appreciable loss in efficiency. In addition to this, gears offer a method of positive indexing that is lacking in other types of drive. I have never seen a belt-driven watch, an electrically-driven camshaft, or a hydraulic link turning the distributor in a piston engine.

design jobs that we do every day, and which we can now do more quickly and more accurately.

Not only do the computers allow the designer to solve problems in a hurry, they also allow him to see a problem in its entirety, so that he can have confidence in his solution. He can change all of the variables to assure himself that he has achieved an optimum design rather than just an acceptable design. He can formulate problems analytically, study maps of solutions to see what the effect of changes will be, and thus analyze in a few moments the results of a lifetime of repetitive calculations.

Another good thing about computers is that they never make mistakes. Well, hardly ever! The novice in the computer field, upon finding that he has obtained a wrong answer, will search his program and, finding no errors, satisfy himself that it must be a machine error, but experience proves this is seldom the case.

All this is not to say that you can take a short problem, feed it to a computer, and get a good answer more quickly than you can do it manually. You have to talk to the computer in a language it understands and tell it every decision it must make. Having listed all of the possible decisions it can make, one then runs through a sample problem and invariably will find that something has been overlooked. The machine has opened up new avenues for error. The computer may work faster than 50 men, but given the wrong set of instructions, it seems to be able to make mistakes faster than 500 men. On the other hand, once a computer has been "debugged" it can be used to feed in new problems of the same type and get correct answers in a matter of minutes.

Gear problems were among the first that we put on our computer, and the gear design people have been constant customers of our computer installation. These problems are not at the top of the list insofar as distribution of total computer time is concerned, but there is one thing peculiar to gear calculations—we don't seem to have a dead file on them. Every program we have ever used has undergone continuous modification and is still being used. Typical of such problems are: external and internal gear data for spur and helical gears, including form factors for bending stress calculations, spur and helical gear meshing characteristics, tooth overlap, clearances, and backlash. We make frequent design and cutting calculations for hypoid gears. We have made mesh analyses of crossed axis helical gears and we have made calculations using the balanced-life concept of gear design.

I will describe in more detail one of these types of gear design that we were able to explore in depth only because we had a computer available to do the repetitive work for us. I refer to the balanced-life concept of gear design in which both gear and pinion are designed to fail simultaneously like

> ". . . the wonderful one-hoss shay,
> That was built in such a logical way,
> It ran a hundred years to a day . . . .
> It went to pieces all at once, ——
> All at once, and nothing first,
> Just as bubbles do when they burst."[1]

The gear design techniques described today must necessarily be limited to a narrow segment of a very large field. Available time alone would necessitate narrowing the choice of subjects and examples to a small fraction of the truck and passenger-car gearing problems which constitute the daily activity of an automotive company. In addition, this writer lays no claim to the technical sophistication required for a detailed view of such a broad field, and

[1] "The Deacon's Masterpiece," by Oliver Wendell Holmes, (*adapted*).

can only present an outlook based on the necessity of accomplishing an overall job in the best and most straightforward manner.

In the design and development of many automotive and truck components, gears are a determining factor; however, the gears themselves are not the primary aim of the research and development effort being carried on. An automotive powerplant, for example, falls directly in this class. Engine development itself is the desired result; but the engine cannot function without gears, so they become a vital part of the design. Therefore, in the interest of reducing the total number of design and subsequent development hours which must be applied to the creation of the powerplant, it becomes extremely desirable to provide gears which can be designed for special application in a minimum of time, which will be as close to optimum size and performance as possible, and which will work the first time and continue to function quietly for the expected life of the entire component.

As a specific example of the type of gearing to be considered in this paper, I could have chosen those in the reciprocating engine, since the design method to be described is just as applicable to those gears as it is to any other type of gearing. However, because of the interest in gas turbines, I have chosen it as a design example to illustrate the balanced-life approach to gear design. This design approach must be used for most gas turbine applications, but is not restricted to high-speed gearing and is equally applicable to the more everyday types of gearing encountered in the automotive industry.

The significant difference between gas turbine and more conventional automotive gear design is that it involves extremes in speed of operation within the same engine. Gas turbine gears are a particular challenge to the gear designer, since admittedly they are more difficult subjects on which to practice his science and his art. The ability, for example, to produce a gas turbine reduction gear with a ratio of 9 or 10/1 in a single step (Fig. 1), to have it work right the first time, and continue to work quietly and efficiently with extended life, is a technical achievement in itself. In addition, during the initial phases of a new engine size determination, the necessity of quickly providing the powerplant designer with a series of gearsets, any one of which may be chosen and each one of which will be equally good, approaches the ideal of engineering design work.

The high-speed gears used in turbines have several characteristics which separate them in design and development complexity from the types of gears ordinarily found in engine design. Generally speaking, they must be more accurately processed, they must have a higher degree of accuracy in their mounting, and the mounts themselves must be more rigid than those usually found necessary for lower

Glossary of Terms

BR = Base radius, in.
CR = Contact ratio
D = Pitch diameter, in.
F = Face width, in.
L = Length of action from pitch point to outside diameter, in.
n = Rpm
N = Number of teeth
N_a = Length of action, in.
PVT = Scoring factor
r = Gear ratio
R = Pitch radius, in.
s = Bending stress, psi
s_c = Compressive stress, psi
T = Torque, lb-in.
T' = Torque at bending design life, lb-in.
X = Tooth form factor, in.
ρ = Radius of curvature at point of contact at top of tooth, in.
ρ' = Radius of curvature at point of contact at bottom of tooth, in.
ϕ = Pressure angle, deg

Note: When subscripts p and g appear, they refer to pinion and gear respectively.

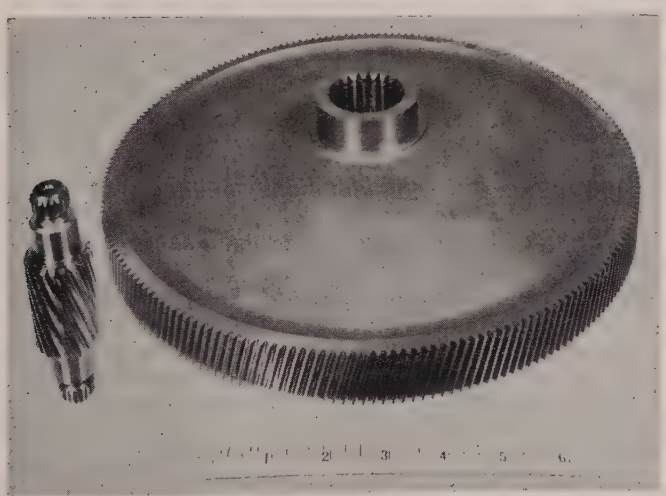


Fig. 1 — Main reduction gear for automotive gas turbine

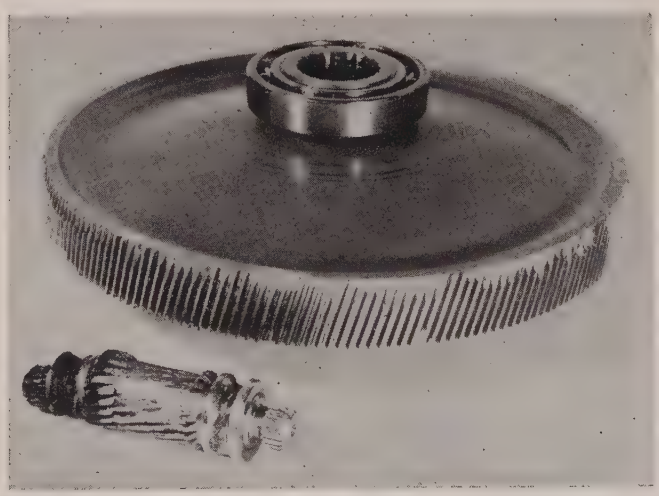


Fig. 2 — Typical failure of gearset operating at high speeds

speed gear trains. Failure to start with these fundamentals is a shortcut to trouble, and when high-speed gears fail, the result is frequently catastrophic (Fig. 2).

High-speed machinery is becoming more familiar to the engineer whose experience has been with engines and motors operating at low speeds. In our turbine, it is possible for a gear mounted on the high-speed output shaft to go through as many as 10 million cycles in about 3 hr. This sort of operation can give a rapid accumulation of fatigue damage in a short time. Gas turbine design today is at the stage where frequent changes in the speed and torque characteristics are made in order to evolve the best powerplant configuration. Where large speed reductions are required, the gear design can be a major factor in the layout of the powerplant. The designer cannot wait for the traditional methods of gear design to give a complete answer before major engine dimensions are set down. The balanced-life design charts allow the selection of gear dimensions quickly, even though the ultimate refinement of the gear design must be finally accomplished by more detailed methods. Working with the preliminary data given by the charts, final dimensions can be fixed by the conventional gear design process or by computer analysis of the problem. In all cases we have found it necessary to make only minor adjustments in engine layout to accommodate the final gear design.

Balanced-Life Gear Design

There are two characteristics which may be related directly to the balanced-life concept of gear design: (1) the adjustment of the thickness of the pinion and gear teeth to compensate for the variation in life between the gear and the pinion, and (2) the adjustment of the outside diameters to give equal scoring factors. It is a desirable goal that the number of cycles to produce fatigue failure be in proportion to the gear ratio, so that both gear and pinion fail at about the same time. There seems to be little justification for having a usable gear destroyed by a pinion which has failed, so in effect the pinion is strengthened at the expense of the gear. Such a gearset has a longer life and no added penalty in size or cost. This approach is necessary to get the most gear capacity into the smallest space.

Balanced-life gears are used in our gas turbines to combine minimum size with maximum capacity. To obtain comparable capacity, the usual gear design practice would give larger pinions. This could lead to the requirement of several stages of reduction gearing to keep the gear train within the envelope of the vehicle outline.

Balanced-life design concepts are not a solution to all gear problems. If gears of standard proportions are available and they are of sufficient size and capacity to do the job, it is unnecessary — and indeed unwise — to spend the extra engineering time and effort to develop a balanced-life design for experimental application. Frequently, the analysis of the problem will show that a pinion of standard proportions will outlive other components, and will define a gearset of reasonable size that is capable of carrying the load. In these instances, it would be unnecessary to use even the most short-cut phases of the balanced-life concept. Further, if size, load capacity, and life are not critical, it is most desirable to economize by using available standard tools. With balanced-life gears, since there is a variation in tooth proportions between the gear and pinion, no two designs are likely to require the same tooling. It also will be found that a different cutter must be used for each piece in a gearset. In addition, balanced-life gears will often have pressure angles that are not usually available in standard cutters, which also precludes the use of standard tools.

Bear in mind that these must be practical gears which are used for varying loads and speeds, and assumptions are necessary in order to take this factor into account. Gears for automobiles and trucks cannot be designed to fit into impossibly simplified vehicles which are expected to run solely at a constant speed of 30 mph for 100,000 miles. Shown in Fig. 3 is a gas turbine accessory drive set which must operate under widely varying conditions of torque and speed.

Gear Failure

An analysis of the methods in which gears fail and how these failures affect the balanced-life design concept must include the following major factors:

Wear
Bending
Scoring
Pitting
Shock

Wear — Some gears wear out. Fortunately, however, high-speed gears are not usually faced with this problem because they must be provided with ample lubrication in order to carry away the heat which is generated by the high sliding velocities.

Bending — Repeated bending of the gear teeth can cause a typical fatigue type of failure. This is the major stress that can be proportioned through the use of the balanced-life design concept. While the load may be within the elastic limit, its repeated application may cause a crack to form at the surface, which reduces the effective area of the tooth and causes further propagation of the crack. If a fatigue crack is found in one tooth, cracks will usually be found in others.

Scoring — Some gears fail by scoring. It might be said that gears which fail in scoring, wear out before they wear in. In high-precision gears of hardened steel, however, there is little wear-in problem. Scoring shows up in what appears to be an abrasive type of wear near the tip or the root of the tooth, and is due to minute welds that take place at points of minor surface roughness which penetrate through the oil film. High-speed gears which fail by scoring will usually do so upon the first application of the critical load and speed.

Experience has shown that an analysis of the design can determine whether the gears will score, since the maximum scoring factor occurs under conditions in which the product of the speed and the square root of torque reaches its maximum value. If the analysis reveals a "no-score" answer, one need not resort to a scoring balance. The calculations

for determining the tendency to score consider the heat generated in sliding from the pitch line to either the beginning or end of tooth action.

Pitting — Gears also fail by pitting. This is another type of fatigue failure in which repeated application of load causes small particles to flake out of the surface of the tooth, leaving small depressions or pits. This is a compressive stress phenomenon, and because the stress areas in the gear and opinion are equal at the point of contract, we do not attempt to balance out life in this regard. However, pitting can usually be avoided by the proper selection of materials and heat-treatment. In most design work, however, steel gears are not likely to fail in pitting if calculations show that the compressive stress is below an acceptable level. Usually, by keeping the compressive stress below 200,000 psi, it is possible to avoid pitting failure and approach infinite life as far as this factor is concerned. It has been our experience in high-speed gears that one is not likely to encounter the compressive stress troubles normally met in more conventional designs. This has been attributed to the extreme rapidity of tooth action, which does not allow the supporting oil film to break down during the time of pure rolling contact at the pitch line. In effect, therefore, the load is applied over a greater area, reducing the compressive stress given by the theoretical calculations.

Shock — Failure because of shock loads may take several forms. A single application of an overload can cause fracture if it exceeds the ultimate strength of the tooth. Repeated application of less severe shock loads may lead to premature fatigue failure in either bending or pitting unless design allowances have been made.

Factors in Gear Life

Some of the conditions leading to failure which the gear designer must consider have been discussed, but what are the factors involved in developing an alleged superior balanced-life gear? The major ones he must work with are:

Gear ratio
Pressure angle
Contact ratio
Scoring factor
Bending stress
Compressive stress

Most of these are old acquaintances and it is only in the way that the balanced-life concept handles them that differences arise.

The variables listed above are functions of both the gear geometry and its environment, and a change in any one variable will usually influence the others. The other well-known variables such as pitch, number of teeth, torque, loading, material, operating conditions, and lubricant are familiar to all gear designers. The methods of handling these are well established, and it should not be necessary to go into them in detail.

Design charts have been prepared for balanced-life gears. Those shown in this paper are based on gear design practice which has been proved for steel gears. If further tests show that the present stress-life data are too conservative, it is necessary only

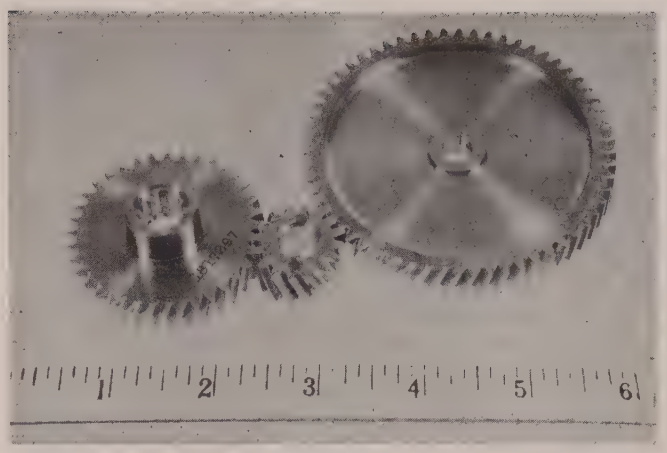


Fig. 3 — Gas turbine accessory gears, high-speed pinion in center

to make minor adjustments to use the new stress cycle charts without obsoleting the major design curves. It is possible to apply these charts to other materials. From the practical viewpoint, however, the effort to get maximum capacity out of minimum package has precluded work on weaker materials; nor has much work been done with fancy alloys, since our gears must be priced for mass-production applications after they have been developed.

The charts used for preliminary balanced-life design vary both the outside diameter and the thickness of the teeth in order to balance the bending fatigue lives and scoring resistances of the mating gears. In addition to this, they assume use of a cutter having a full radius tip in generating the minimum active dedendum. This is not allowable under all conditions, of course, especially where there is a bored hole through gears of relatively small diameter.

An assumption was made that the backlash is equal to 0.04 divided by the diametral pitch. This was chosen as a representative value for this type of gear. Admittedly, this is an empirical assumption which might not hold true if the load did not reverse. For reversing loads, however, it is sufficient to allow for net changes in length and stack-up tolerances. The variation in tooth thickness necessary to give other values of backlash would be of negligible magnitude.

Gear Ratio — Even in preliminary designs, the gear ratio is pretty well fixed by the time the gear designer is called in to express an opinion. Normally, one would like to have as few gears as possible, and we have successfully made one-stage gear reductions in our turbine in excess of 9.5/1. In the past, gear designers frequently added one tooth to the gear to avoid integral gear ratios. The hunting tooth, as it was called, was added to keep the same teeth from coming into contact repeatedly. It avoided some problems of noise and allowed the teeth to wear in without causing too much damage. Working with hardened gears of extreme precision, one need no longer be concerned with the problem of wearing in, and it is not felt that the hunting tooth is required.

Noise enters into the selection of gear ratio and it is most desirable to avoid mesh frequencies high in the audible range. In gas turbines, this usually

requires taking a large reduction in the first step when it is necessary to have more than one stage of reduction. This also tends to reduce power losses in bearings due to running at high speed.

Pressure Angle — The next step is to determine the pressure angle. This is defined as the angle between the tangent to the tooth profile at a pitch circle and the perpendicular to the pitch surface at the same point. Selection of the proper pressure angle is important because of the influence it has on the tooth as a loaded beam, as illustrated in Fig. 4. Large pressure angles tend to reduce both bending stress and compressive stress on the teeth, and they allow the manufacture of pinions with fewer teeth, without experiencing difficulties due to undercut. There is a limit, however. If a design results in pointed teeth, the gears will be more subject to accidental damage, and also cannot be hardened without serious risk.

The charts shown have been prepared for pressure angles of 14½, 20, and 25 deg. Other things being equal, it is generally accepted that a lower pressure angle will result in a lower capacity, quieter gear. In the past, a 14½-deg pressure angle was commonly used, but with the introduction of more accurate gear tools and the desire for smaller sizes, a 20-deg pressure angle has come into common use. In specialized applications the designer may want to use 25-deg or higher pressure angles, but these should be restricted to designs where the major consideration is capacity.

Contact Ratio — The gear design charts shown have been prepared for a contact ratio of 1.5. This is a good compromise between the lower values which are usually associated with increased bending strength and the higher values which usually result in quieter gears, less likely to score. Contact ratio variations can be influenced by production tolerances, rigidity of mounting, and the number of teeth in the gears and pinions. In gears with smaller numbers of teeth, the contact ratio is more sensitive to mounting errors.

Scoring Factor — Some years ago Straub[2] defined scoring factor as a function of the total heat generated during the cycle of sliding from the pitch line to either the beginning or the end of tooth action. With the gears he tested, it was found that the maximum allowable design scoring factor, using mineral oil, was 1½ million. Our tests have been in general agreement with these findings, though undoubtedly there is room for additional testing to define this limit more accurately for high-speed gears.

Bending Stress — Bending of gear teeth is the major factor that the designer of balanced-life gears can take into consideration. We have found the greatest success with the methods of analysis of Almen and Straub.[3] We know of no theoretical approach that gives better results, even though today it is possible to use computers for rigorous analysis of stresses with little concern for normal computing time as a major factor.

Compressive Stress — Compressive stress is calculated by the classical method of Hertz.[4] Wickenden, *et al.*[5] performed tests to determine the stress-life relationship and to establish the stress below which the gear life is indefinitely long. These curves have been used together with one obtained from "Modern Methods of Gear Manufacture"[6] for the compressive stress analysis. However, the hypothesized influence of a continuous lubricant film previously mentioned allows operation which approaches infinite life at higher stress values.

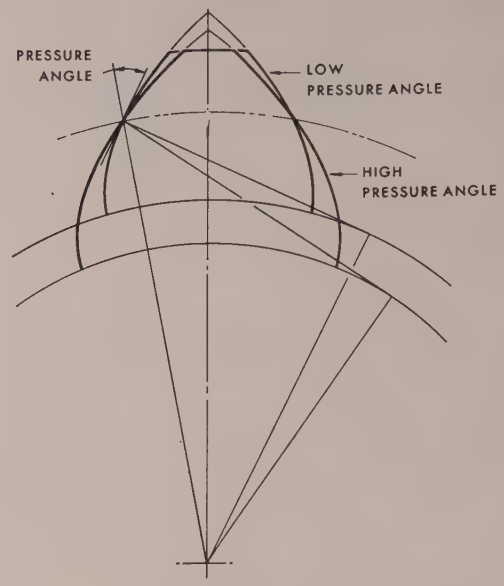


Fig. 4 — Relationship of pressure angle to tooth contour

With regard to the equations used and the method of approach, the maximum allowable outside diameter as limited by the desired scoring factor is first defined. Using Straub's equation for scoring factor, begin with these two equations:

$$PVT_p = 19.984 \frac{(1+r)^{\frac{3}{2}}}{r} \sqrt{\frac{\sin\phi}{CR\cos\phi}} \frac{L_p^2}{\sqrt{\rho_p \rho'_g}} \sqrt{\frac{T_p}{F}}\, n_p \qquad (1)$$

$$PVT_g = 19.984 \frac{(1+r)^{\frac{3}{2}}}{r} \sqrt{\frac{\sin\phi}{CR\cos\phi}} \frac{L_g^2}{\sqrt{\rho'_p \rho_g}} \sqrt{\frac{T_p}{F}}\, n_p \qquad (2)$$

Equating the scoring factor at the outside diameter of the gear to the scoring factor at the outside diameter of the pinion, it is possible to cancel out all but the following factors:

$$\frac{L_p^2}{\sqrt{\rho_p \rho'_g}} = \frac{L_g^2}{\sqrt{\rho'_p \rho_g}} \qquad (3)$$

In these equations $(L_p + L_g)$ is equal to the active length of contact between the two gears. (See Fig. 5.) If a value is assumed for either L_p or L_g, it is possible to solve for all other factors in Eq. 3, and by inspection, make a more accurate second assumption for a value of L_p. Interpolation between the two values of L_p will yield a result which defines a

[2] "Shotpeening as Factor in Design of Gears," by J. C. Straub. *Mechanical Engineering*, Vol. 73, July, 1951, pp. 565–569.

[3] Factors Influencing the Durability of Automobile Transmission Gears, Part 2," by J. O. Almen and J. C. Straub. *Automotive Industries*, Vol. 77, Oct. 9, 1937.

[4] "Formulas for Stress and Strain," by R. J. Roark, Case 5, p. 226. Pub. by McGraw-Hill Book Co., Inc., New York, 1943.

[5] "Evaluating Surface Durability of Gears," by T. H. Wickenden, G. R. Brophy, and A. J. Miller. *Machine Design*, Vol. 18, July, 1946, pp. 142–146.

[6] "Modern Methods of Gear Manufacture." National Broach and Machine Co., Detroit, 1950, third edition.

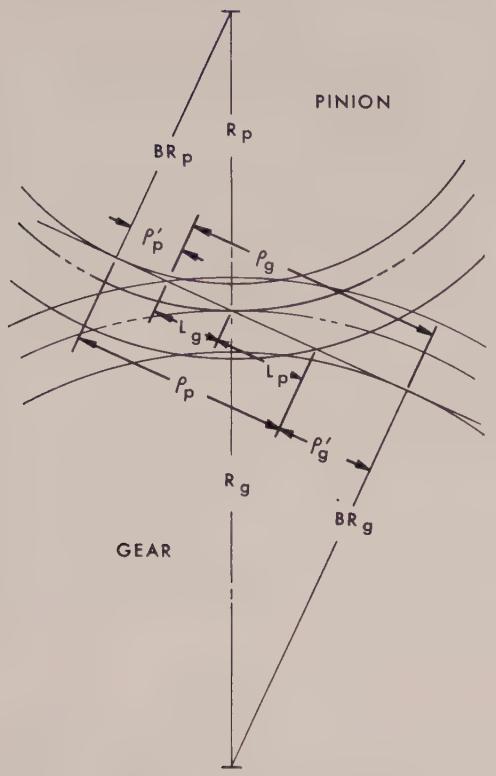


Fig. 5 — Line of action diagram

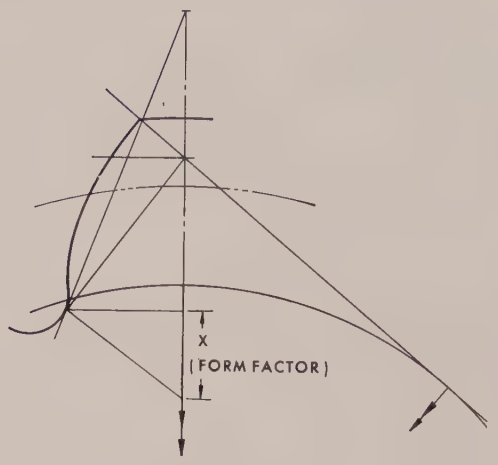


Fig. 6 — Tooth form factor

pair of gears with equal scoring factors. Returning now to Eqs. 1 or 2, the proper face width for the gears can be determined. By such successive trials, the scoring life can be balanced.

Next, the bending stresses are balanced. A bending stress for the pinion must be found which will give a life in cycles which is related to that of the gear as the inverse of the gear ratio, assuming that infinite life of both of these items has not already been attained. In other words, a condition is specified in which both gear and pinion will reach their endurance limit in bending at the same time. Examination will show that all of the major gear dimensions have been determined at this point except the tooth thickness. One of the factors which has been established is the circular pitch, which is equal to the thickness of the tooth of the pinion plus the thickness of the tooth of the gear plus the backlash. Since a value for backlash has been assumed, this figure can be subtracted from the circular pitch to obtain the sum of the two tooth thicknesses. Successive values for tooth thickness are now assumed until one that gives equal bending stress fatigue life is obtained. Here is where there is no substitute for experience in the field of gear design. Ignoring experience, however, suppose for the moment it is assumed that the two teeth are of the same thickness. A tooth form factor (X factor) can be determined graphically from a gear layout (Fig. 6). In practice, however, it is easier to do this by analytical methods on a computer. Before proceeding, it should be verified that the tooth thickness is not such that is will result in a pointed tooth. It is now possible to go to Eq. 4 and calculate the bending stress values for the pinion:[7]

$$s_p = \frac{3\pi T_p}{F N_a N_p X} \tag{4}$$

A similar equation will give the stress in the gear.

Another assumption may now be made of the relative proportion of the tooth thicknesses and the stress calculated for this condition. A graphic plot is made of the ratio of the tooth form factors of the gear and pinion versus the tooth thickness of the pinion. From the equation:

$$\left(\frac{X_p}{X_g}\right)^{5.126} = r\left(\frac{F_g}{F_p}\right)^{5.126} \tag{5}$$

which gives the optimum ratio of tooth form factors, the value of the optimum pinion tooth thickness can be determined from the plot.

The gearset now has both balanced scoring life and bending fatigue life. One more check remains before the design is complete — the evaluation of compressive stress. As was indicated, this computation is based on the Hertz equation for two cylinders in contact along a common element.[8] From this equation the compressive stress at the pitch line of steel gears can be derived, assuming that one tooth is in contact. It may be that the compressive stress is too high to give satisfactory operation. A safe level for the desired life can be determined by reference to the following equation:

$$s_c = 2290\sqrt{\frac{T_p}{R_p^2 \cos\phi \sin\phi F}\left(1+\frac{1}{r}\right)} \tag{6}$$

If this determination indicates an unsatisfactory stress level, it may still be possible to avoid the necessity for recalculating the entire design. This requires the use of another face width which will reduce the stress to a satisfactory level and still

[7] Almen and Straub, *op. cit.*
[8] Roark, *loc. cit.*

have a gear of reasonable proportions. If the maximum usable face width still results in a compressive stress which is too high, it is necessary to go back to the beginning and assume a new value which will increase the radii of the equivalent cylinders at the pitch point. This can be done by increasing the center distance of the two gears or by increasing the pressure angle and maintaining the same center distance.

Computer Programming

An examination of the basic equations that have been balanced indicates that they are made up of two types of factors. The first group consists of the physical characteristics of the gears themselves. The second covers other characteristics having to do with conditions under which the gears will operate. These equations can be rewritten to equate the gear geometric characteristics to the operating characteristics in each case. When this had been done, a program was prepared for a digital computer, which would work through the complete calculations for a set of balanced-life gears. In effect, the computer was given the description of the gear, its diametral pitch, number of teeth, contact ratio, pressure angle, and gear ratio and then was instructed to solve for the nongeometric values for both a balanced scoring life and a balanced bending stress life. A range of values of each of the geometric characteristics was selected and the computer ground away, busily solving problems with which the gear designers hadn't yet been confronted. The tables that have come out of this program are, in effect, a list of presolved balanced-life calculations. All of the calculations that were done on the computer were the same as those that would have been done in a manual program. The computer gave values in terms of a constant so that each of these factors could be plotted on a grid in which all the nongeometric functions of bending stress are in the ordinate and all of the nongeometric functions of the scoring are in the abscissa.

Only one complication remained at this point. The data required a three-dimensional chart, and three-dimensional charts are hard to get on two-dimensional paper. To avoid this extra dimension, the face width factor was taken out of the geometric functions and placed on the other side of the equation, so that it appeared as part of the coordinate system. The charts were then constructed for a face width of 1 in. They can then be easily entered for this single value of face width, and the value thus obtained converted to any other reasonable face width by the application of a simple nomograph.

The complete program is such that the computer can always be given instructions to solve any balanced-life gear problem. There are two drawbacks here, however: (1) Computer time may not be immediately available when it is needed to make quick analyses; (2) the computer cannot separate unreasonable solutions which are due to conditions other than those in the basic gear design equations. For a very versatile program it has been better to make the machine calculations and inspect them later for anomalies, rather than to list all possible pitfalls, together with methods of avoiding them.

A similar approach was used for obtaining compressive stress charts. Once again, the face width was moved from the geometric side of the equation in order to simplify plotting.

Figs. 9–14 are a family of charts in which a number of spur gear problems were solved using the balanced-life concept.

What of the future of gear designs as influenced by computers? Today, we are using computers more and more, and still we need more people in our gear design departments. We have new types of gears being added constantly and these problems can be analyzed better on the computer. We can extend or refine the analytical gear design concepts into nearly all applications, and arrive at the best gear for each use without relying on today's handbook assumptions. In addition to this, there are bound to be new things coming up for computers and for gear theory that we don't even know about today. It is likely that the computers will evolve new methods of analyzing gear stresses and will accelerate the progress of the field more than we can anticipate. One thing is certain; today's work in computers will relieve the gear designer of much of the tedium of his job. He will have more time to spend in analytical work.

Typical Problem

A typical gear design problem will be examined to show how some of these charts might be used.

For a preliminary design, the engineer will usually be furnished with the gear ratio, approximate center distance, and the conditions under which the gear must operate. For example, assume that specifications which include the following have been established:

Ratio	4/1 reduction
Approximate Center Distance	5 in.
Operating Conditions	No extremes of temperature Medium viscosity mineral base oil Negligible shock loading Torque-speed curve

From this and other information in the general specifications, it is possible to determine the design life for bending and compressive stresses. The operating conditions under which gear life will be specified must now be selected. In the example, this was estimated by the designer to be two million cycles at a higher-than-normal sustained operating torque of 42.3 lb-ft. Based upon this, an allowable bending stress of 52,000 psi was determined for a carburized nonpeened gear with a hardness of R_c 60. This information is taken from the bending stress-cycle chart (Fig. 7), a fatigue chart based on experimental data.

The next step is to find, from the torque-speed curve in the specifications, the conditions at which maximum tendency to score will occur. This is the point where a maximum value is reached for the factor, $\sqrt{T_p}\ n_p$—in this case 219 lb-in. of torque at 45,000 rpm. The scoring factor at this point on the curve should not exceed 1,500,000. Gear design charts can now be selected. (See Figs. 9 through 14). Because a pressure angle of 20 deg would be satisfactory for this application, this narrows the selection to a single pair of charts (Figs. 11 and 12) with

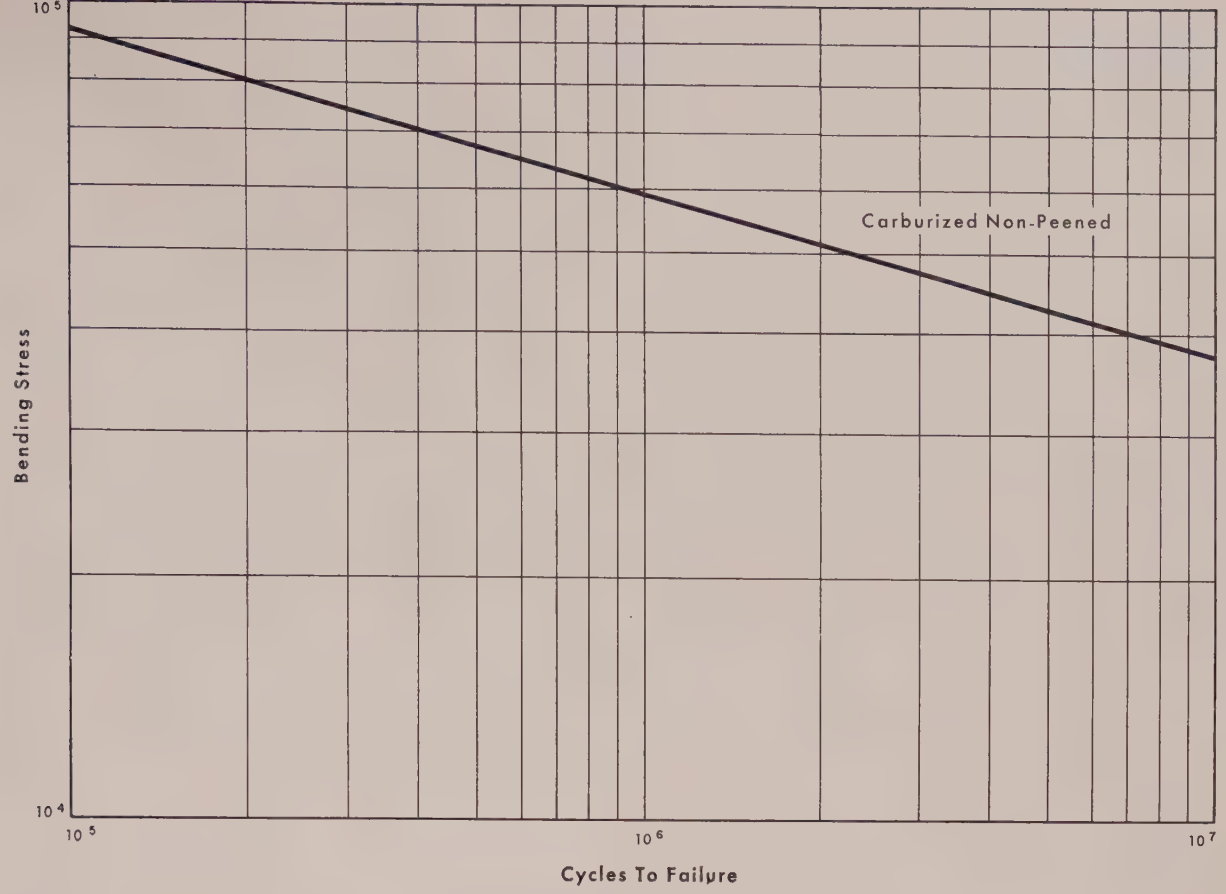


Fig. 7 — (from "Durability of Automotive Gears," by Almen and Straub)

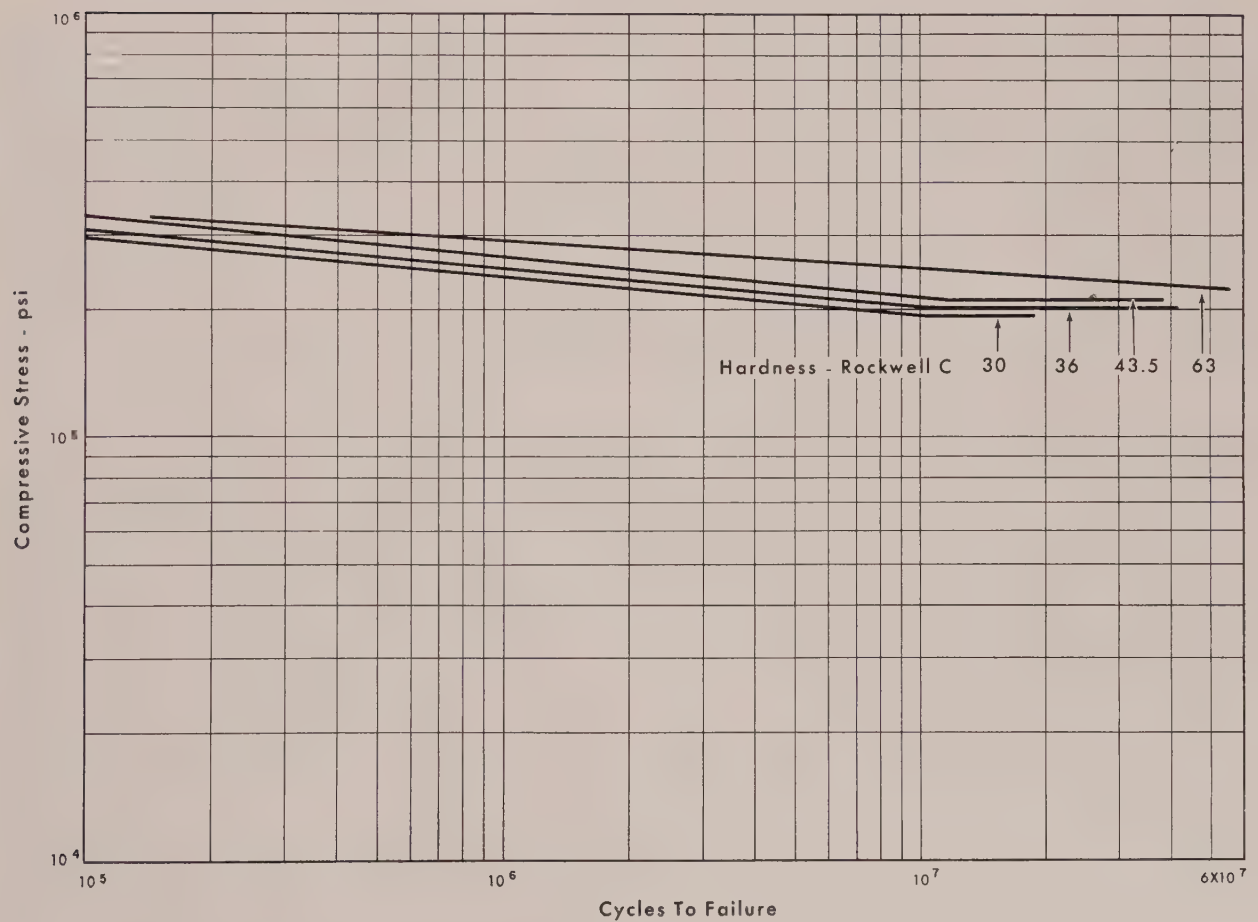


Fig. 8 — (from "Evaluating Surface Durability of Gears," by Wickendon, et al, and "Modern Methods of Gear Manufacture")

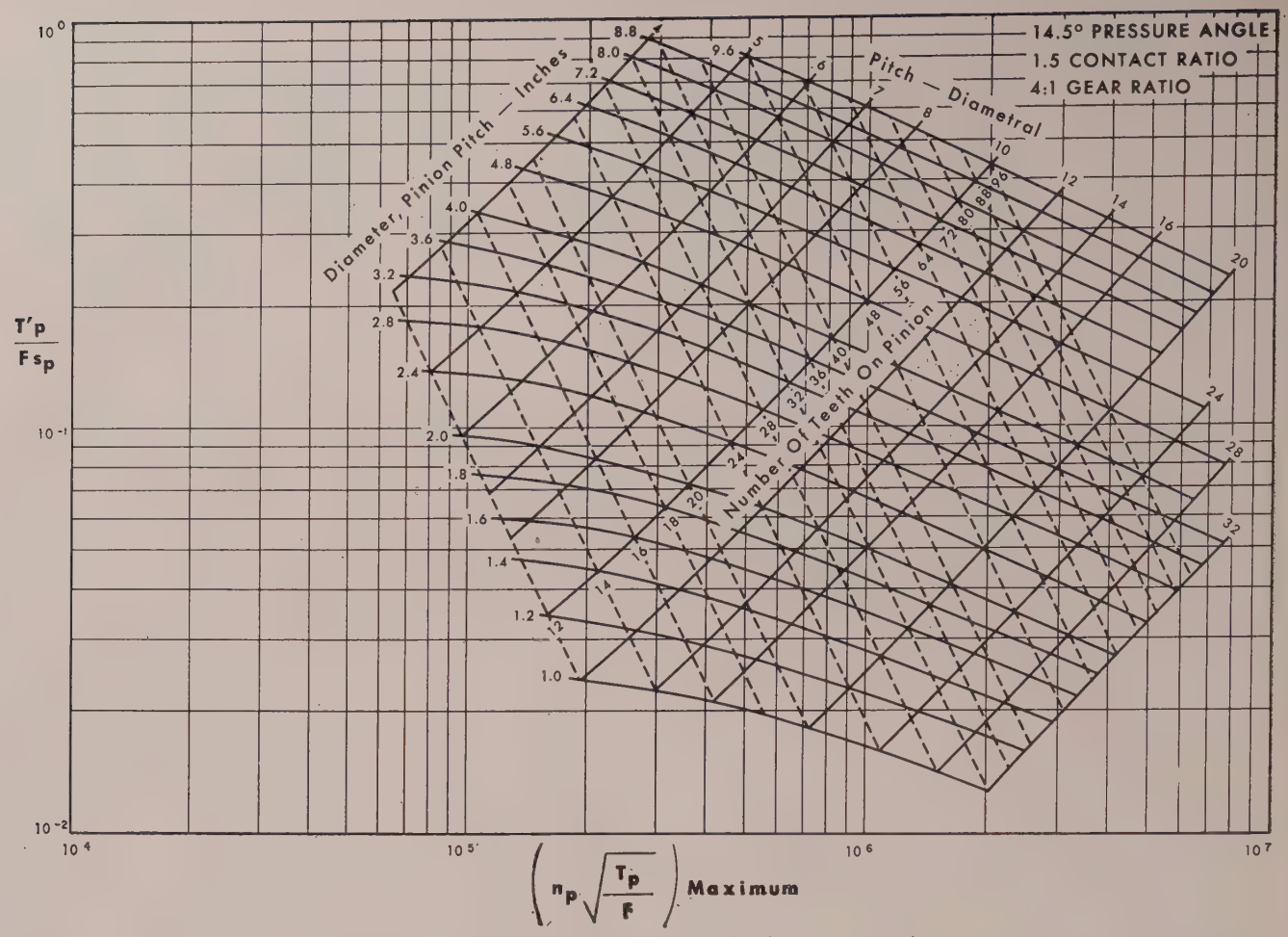


Fig. 9 — Balanced-life gear design chart, 14.5 deg pressure angle

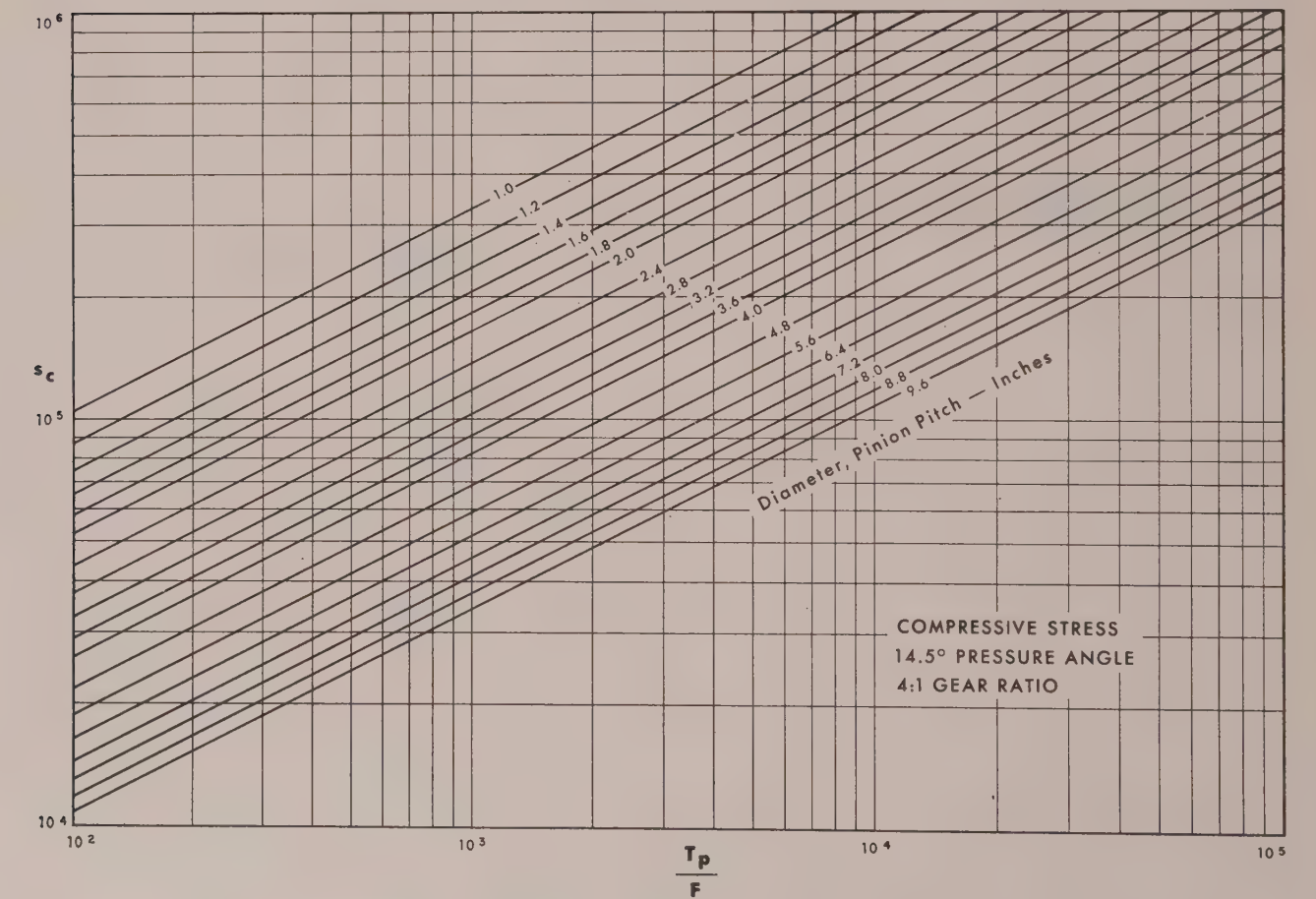


Fig. 10 — Compressive stress chart, 14.5 deg pressure angle

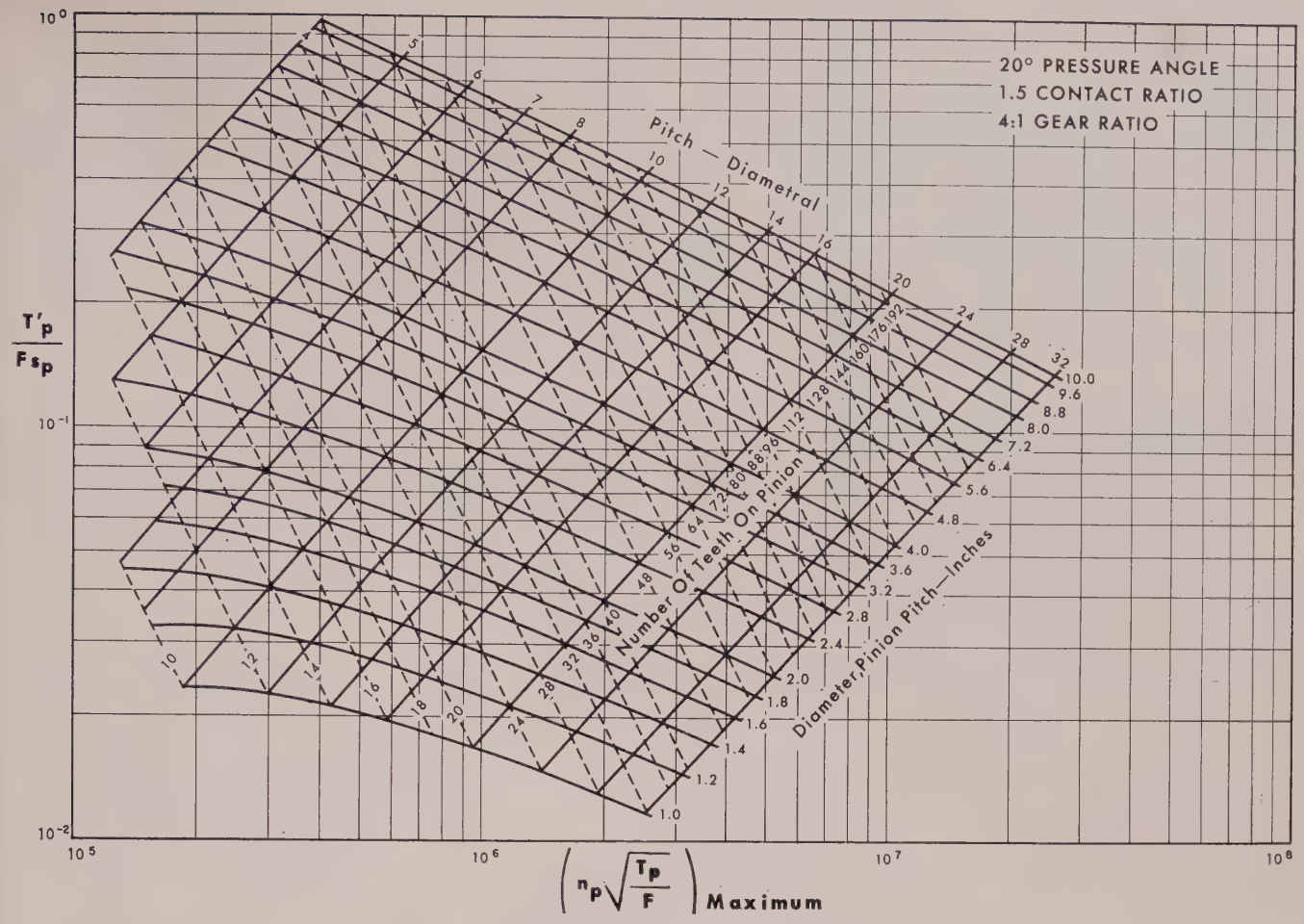


Fig. 11 — Balanced-life gear design chart, 20 deg pressure angle

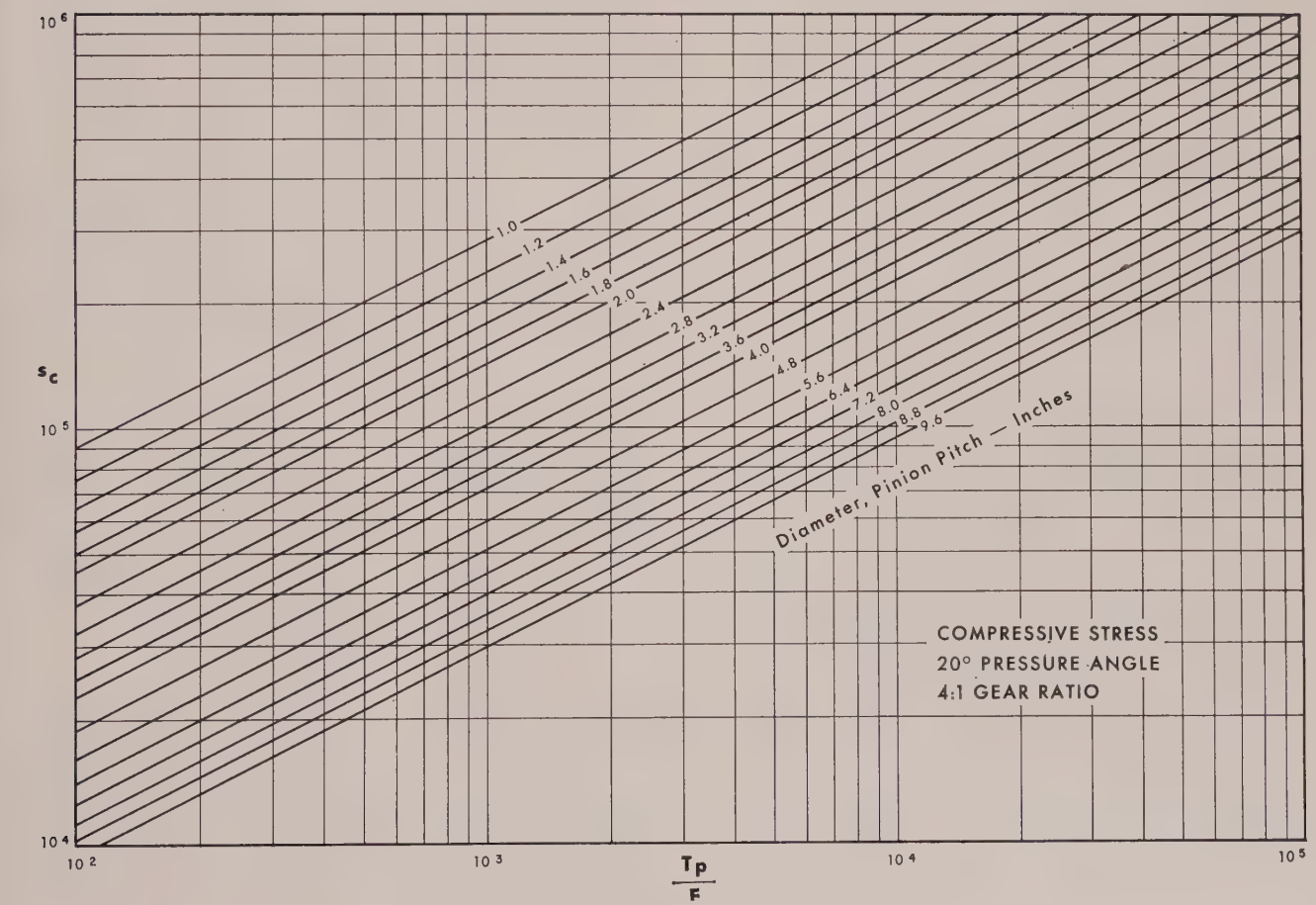


Fig. 12 — Compressive stress chart, 20 deg pressure angle

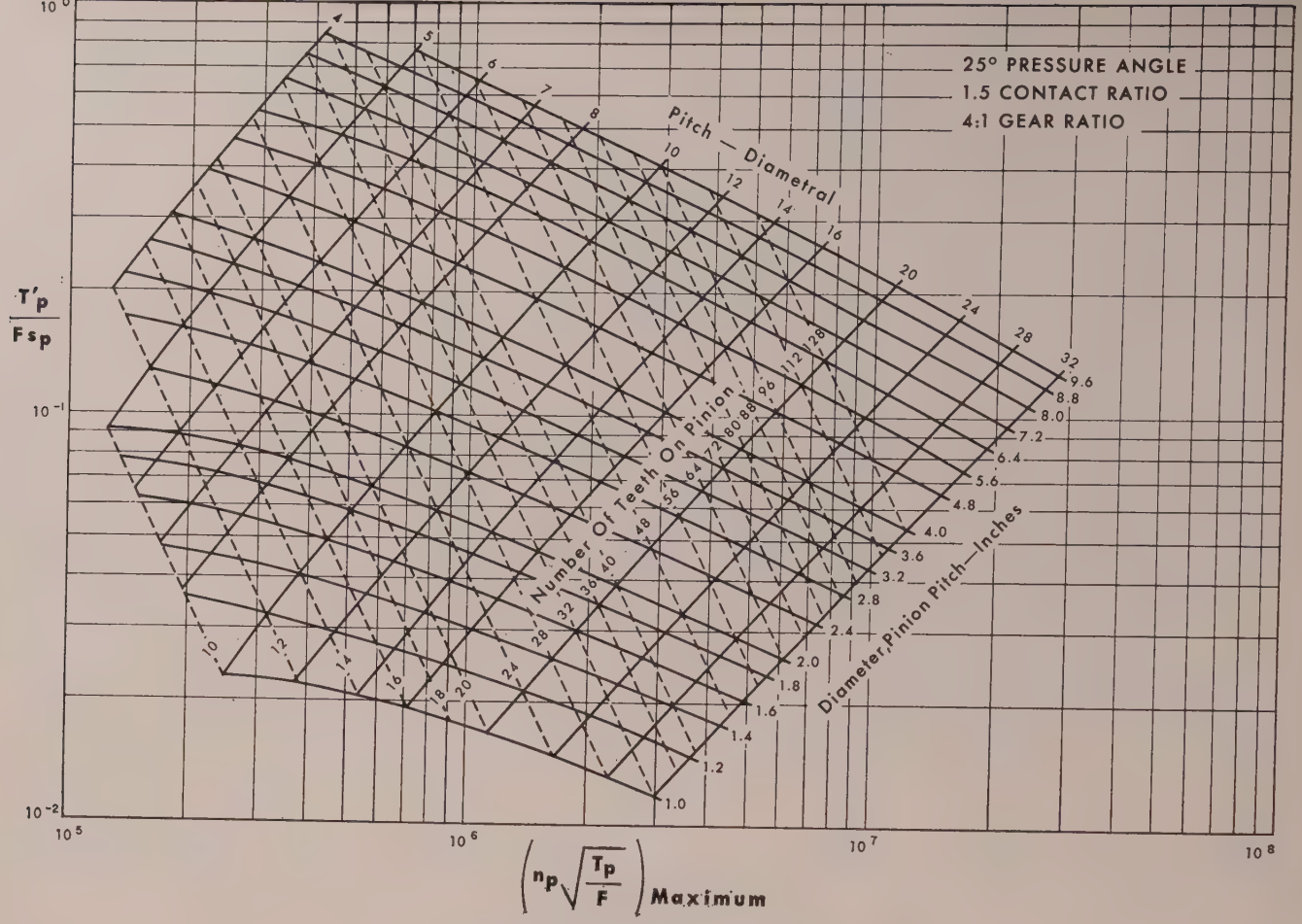


Fig. 13 — Balanced-life gear design chart, 25 deg pressure angle

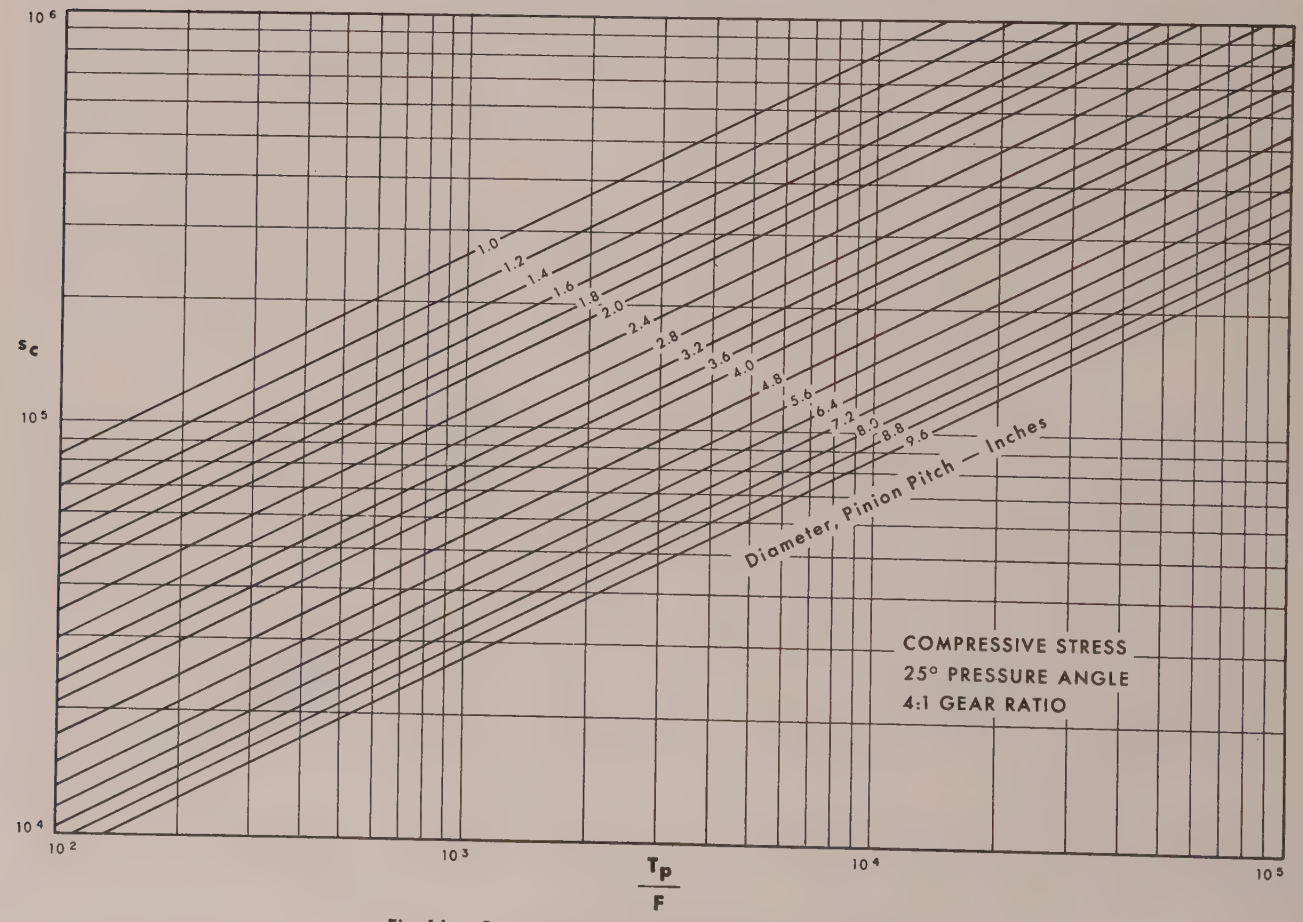


Fig. 14 — Compressive stress chart, 25 deg pressure angle

20 deg pressure angle, 4/1 reduction, and 1.5 contact ratio.

Values of ordinate and abscissa can now be calculated in order to enter the gear design chart. For gears of this type, a tentative value for face width of 0.5 in. seems reasonable. This initial assumption is not too critical because it can be easily varied, as you will see.

For the ordinate:

$$\frac{T'_p}{Fs_p} = \frac{507}{0.5(52{,}000)}$$
$$= 1.95 \times 10^{-2}$$

Any point on the grid having an ordinate value of 1.95×10^{-2} or above will meet or exceed the requirements for these conditions of bending.

For the abscissa:

$$n_p \sqrt{\frac{T_p}{F}} \quad \frac{\text{Design Scoring Factor}}{10^6} = 45{,}000 \sqrt{\frac{219}{0.5}} (1.5)$$
$$= 1.41 \times 10^6$$

Any value on the chart which is to the right of a vertical line drawn through 1.41×10^6 will satisfy the requirements for scoring factor.

For these conditions, a pinion of 0.5 in. face width described at the intersection of the two lines, has a pitch diameter of about 1.2 in and a diametral pitch of approximately 22.

(Experience indicates that a gear of this diametral pitch would present problems in manufacture. It might crack in heat-treating and the grinding wheel needed would be so fine as to create problems in tool life.)

From the required center distance, the proportions of pitch diameters can be established from the equations:

$$CD = \frac{D_p + D_g}{2}$$

and:

$$D_g = rD_p$$

which yield a value of: $D_p = 2$ in.

This value shows that the pinion pitch diameter previously obtained from the chart does not meet all of the requirements. Knowing that the pinion pitch diameter should be 2 in., the face width nomograph (Fig. 15) can be used to correct this situation. By aligning it properly and placing it on the chart so that the point, $F = 0.5$ in., coincides with the point previously plotted for $F = 0.5$ in. it is possible to read a new value for face width where the nomograph scale intersects the 2-in. pitch diameter curve. This allows the designer to obtain an optimum configuration which meets center distance requirements. From the chart at this point, the following characteristic of the pinion may be read:

Pinion pitch diameter	2	in.
Diametrical pitch	20	
Number of teeth	40	
Face width	0.26	in.

Using the companion compressive stress chart (Fig. 12) proper values may be substituted, and a value for the abscissa determined.

Locating this value on the abscissa, a vertical line is drawn which intersects the line for a pinion pitch diameter of 2 in. From this intersection, a value of compressive stress of 200,000 psi is read on the ordinate.

Referring to the compressive stress-cycle chart, (Fig. 8) it is seen that the life will be in excess of the two million cycles originally specified and pitting will not be encountered.

Through the use of conventional gear formulas, dimensions of the mating gear are established.

Experience has shown that, for the purposes of preliminary layout, it is reasonable to assume a value for the addendum of the pinion and gear equal to the reciprocal of the diametrical pitch in order to obtain complete dimensions for a layout.

This completes the preliminary design of the gearset. Through the use of the equations used in the construction of the charts, the gear designer may determine the detail dimensions required for the final gear design.

Acknowledgment

As is usually the case in any technical effort, this paper does not reflect the work of a single person, but rather the efforts of several people who have been working on a common problem for a number of years. The early concepts of the balanced-life gear charts and the application of the program to computer analysis were largely the work of J. M. Scofield. D. L. Outen, who was also involved in the early work, has carried it on and expanded it to the present form. The computer programs were prepared with the assistance of J. Rogoff. I am also indebted to A. R. White, E. L. Jones, and J. W. Furlong for their valuable contributions in the preparation and review of the paper.

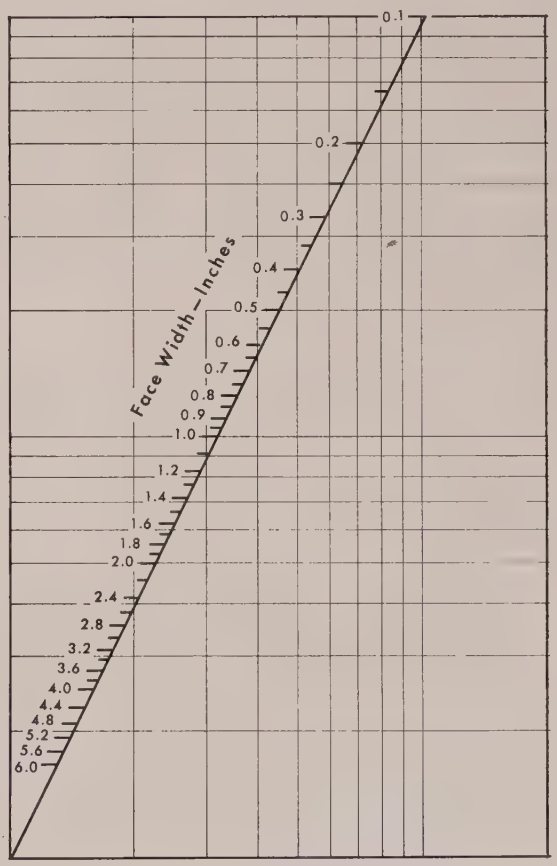


Fig. 15 — Nomograph for various face widths

design procedure for VIBRATION on Nonrigid

Sheldon Rubin
Hughes Aircraft Co.

PAPER DESCRIBES a procedure for the dynamical design of linear vibration isolators to protect a rigid equipment from the sinusoidal motions of a nonrigid supporting structure. A single degree of translational freedom of the equipment is assumed.

Three types of data are required: (1) The equipment mass and its so-called fragility curve, (2) the mechanical admittance of the supporting structure and its amplitude of vibration when it supports no load, and (3) information to establish a minimum stiffness of the isolators.

The procedure yields a region of permissible natural frequency and fraction of critical damping combinations for the isolator-mass system. The specification for the isolator dynamical requirements, including allowable tolerances, is derived from this region.*

THE PROTECTION of airborne fragile equipment from shock and vibrational motions of a nonrigid supporting structure constitutes the activity of the G-5 Aircraft and Missile Shock and Vibration Committee of the Society of Automotive Engineers under the chairmanship of C. T. Molloy. The committee intends to consider the isolation of fragile equipment from sinusoidal, random, and shock excitations, and to produce an SAE Recommended Practice document for the solution of such problems.

Most of the activity to date has been devoted to the generation of an unambiguous step-by-step design procedure which takes into account the behavior of the supporting structure as a finite-impedance mechanical source of excitation.[1] The procedure must be so formulated that it can be applied by a design engineer having only the use of a slide rule and/or desk computor. It is intended to treat center of gravity suspensions and base mounting suspensions (or dynamically equivalent suspensions), since these types form the major portion of vibration isolation applications. A straightforward procedure will provide design engineers with a useful tool for selecting a proper suspension. In addition, such a procedure will serve to define clearly the parameters of the problem, and it is hoped that greater efforts toward the accumulation of the data necessary for the realistic solution of vibration isolation problems will result.

Solving a problem of this complexity is best accomplished in stages. As a first step it was thought desirable to treat the most straightforward case, which could then form a basis for the treatment of the other cases. In line with this conception, the particular set of conditions presented in the next section was chosen for initial consideration. The committee solicits comments and suggestions on the procedure and possible extensions of it.

The Problem

A procedure is to be written for the dynamical design of fragile equipment installations with regard to sinusoidal motion of a nonrigid supporting structure.[3] The procedure is subject to assumptions and limitations as follows:

1. For purposes of formulating the equations of

ISOLATION

Supporting Structures

motion it is assumed that the equipment is a rigid body. More precisely, it is supposed that the input motion to the equipment is not significantly altered by any nonrigid behavior of the equipment. It is acknowledged that resonances most certainly occur within the equipment, but it is assumed that the mass associated with those resonances is relatively small so that their existence does not significantly influence the forces which exist at the equipment support points.

2. It is assumed that the major effects in a given direction of motion can be assessed by considering pure translation of the equipment in the given direction only. This consideration requires that the motions of the equipment support points are nearly enough alike so they do not induce significant rotations of the equipment. In addition, vibration isolation systems are limited to center-of-gravity suspensions. Thus, each of three mutually perpendicular directions can be treated individually.

3. It is assumed that each of the vibration isolators may be described dynamically by a mechanical parallel combination of a linear massless spring and a viscous damper.

4. It is assumed that the susceptibility to failure of the equipment can be described by a single curve of sinusoidal amplitude versus frequency, termed a fragility curve. This curve is the result of single frequency vibration tests on the equipment. The fragility amplitude at each frequency is determined so that there is small probability that a lower sinusoidal input will result in equipment failure. It is assumed, also, that equipment failures occur as a result of the excitation of resonances within the equipment, and that only one of a number of simultaneous sinusoidal inputs provides the predominant excitation at each resonance. Hence, an input having a discrete frequency spectrum can be handled as a noninteracting group of sinusoids which are the components of the input.

Basic Mathematical Relationships

A number of mathematical relationships required in the establishment of a design procedure are given in the following sections.

Equations of Motion — In order to determine the influence of a nonrigid supporting structure, it is useful to consider the supporting structure as a mechanical source of vibrational motion.[4] As shown in Fig. 1A, the source is viewed as a combination of an ideal vibration generator and a linear structure. The performance of such a source at any frequency,

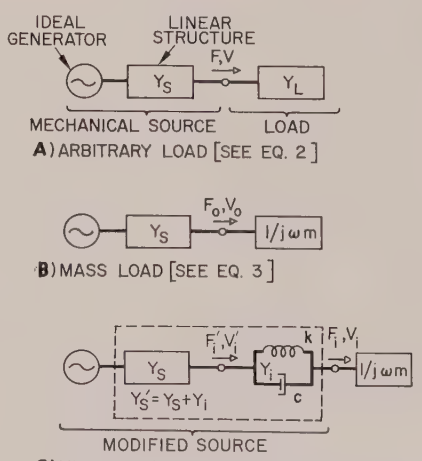


Fig. 1 — Representation of supporting structure as mechanical source and various load conditions

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 7, 1959.

[1] The effect of purely resistive or reactive impedances of the supporting structure on vibration isolation has been investigated by Sykes (Footnote 2). In the present paper the supporting structure is considered to be composed of resonant members.

[2] "Effects of Machine and Foundation Resilience and of Wave Propagation on Isolation Provided by Vibration Mounts," by A. O. Sykes. SAE Transactions, Vol. 66, 1958, pp. 532–548.

[3] Nondynamical considerations are excluded from the present discussion.

[4] "Four-pole Parameters in Vibration Analysis," by C. T. Molloy. Colloquium on Mechanical Impedance Methods for Vibration Analysis, R. Plunkett, ed. ASME, 1958.

ω, may be expressed as

$$F(\omega) = Z_s(\omega)[V_s(\omega) - V(\omega)] \quad (1)$$

where:

F = Force delivered by source to load
V = Velocity of output terminal of source
V_s = Velocity which would exist at output terminal under zero load condition, that is, when $F = 0$
Z_s = Mechanical impedance that would be measured at output terminal with no load present

V_s and Z_s are assumed to be fixed at each frequency, while F and V vary with the character of the load in accordance with Eq. 1. An alternative expression is:

$$Y_s(\omega)F(\omega) = V_s(\omega) - V(\omega) \quad (2)$$

where:

Y_s = Mechanical admittance[5] (reciprocal of impedance) that would be measured at the output terminal with no load present

All quantities in the above equations are complex amplitudes of sinusoidal quantities, and follow the usual convention. The equations apply to each frequency at which motion occurs.

For the problem at hand, it is necessary to consider two different load conditions on the source. The first is a mass load (Fig. 1B), and the second is a vibration isolator interposed between the source and a mass (Fig. 1C). Terming the output velocity for the mass load V_0 and the output force F_0 it follows that $F_0\omega = j\omega m V_0(\omega)$. Substituting this equality in Eq. 2 yields the result:

$$\frac{V_0(\omega)}{V_s(\omega)} = \frac{1}{1 + j\omega m Y_s(\omega)} \quad (3)$$

In the isolator-mass configuration, the isolator consists of a linear spring k_i and a viscous damper c_i. The transfer admittance of the isolator is given by:

$$Y_i(\omega) = \frac{V_i(\omega) - V_i'(\omega)}{F_i(\omega) - F_i'(\omega)} = \frac{1}{c_i + (k_i/j\omega)} \quad (4)$$

It can be shown readily that the system in Fig. 1C behaves as if the mass load were driven by a modified source, the performance equation of which is:

$$[Y_s(\omega) + Y_i(\omega)]F_i(\omega) = V_s(\omega) - V_i(\omega) \quad (5)$$

where:

$Y_s(\omega)$ = Admittance of original source
$V_s(\omega)$ = Free output of original source

Comparing Eqs. 5 and 2 shows that the presence of the isolator at the output of the source has the effect of producing a modified source, the admittance of which is the sum of the source and isolator admittances, and the free output is the same as that of the original source. Therefore, the ratio of the velocity of the mass to the free source velocity, which follows directly from Eq. 3, is:

$$\frac{V_i(\omega)}{V_s(\omega)} = \frac{1}{1 + j\omega m[Y_s(\omega) + Y_i(\omega)]} \quad (6)$$

Generalized Transmissibilities — Dimensionless ratios, termed generalized transmissibilities, provide substantial simplification of the design procedure. The generalized transmissibility function, $H(\omega)$, is defined to be the absolute value of the ratio of the velocity of the load mass to the free velocity of the source,[6] that is:

$$H_0(\omega) = \left|\frac{V_0(\omega)}{V_s(\omega)}\right| \quad (7)$$

when the mass is attached directly to the source (Fig. 1B), and:

$$H_i(\omega) = \left|\frac{V_i(\omega)}{V_s(\omega)}\right| \quad (8)$$

when an isolator is used (Fig. 1C). The H function is merely an expression of the magnitude of the mass velocity as a multiple of the magnitude of the free velocity. (In the design procedure it is necessary to consider a number of different isolators. In order to clearly distinguish the particular isolator being considered, the subscript i is replaced by an integer, i = 1,2,3 . . . , which then serves to distinguish a particular isolator design.)

Eqs. 3 and 6 may be written in terms of H and the real and imaginary parts of the admittances to give:

$$H_0(\omega) = (\{1 - m\omega B_s(\omega)\}^2 + \{m\omega A_s(\omega)\}^2)^{-1/2} \quad (3')$$

and:

$$H_i(\omega) = (\{1 - m\omega[B_s(\omega) + B_i(\omega)]\}^2 + \{m\omega[A_s(\omega) + A_i(\omega)]\}^2)^{-1/2} \quad (6')$$

where:

$$Y_s(\omega) = A_s(\omega) + jB_s(\omega);$$
$$Y_i(\omega) = A_i(\omega) + jB_i(\omega) \quad (9)$$

It may be seen that any procedure for solving Eq. 3′ can also be used to solve Eq. 6′. It is necessary only to modify the source admittance by the addition of the isolator admittance.

It is convenient to express the admittance of the isolator in terms of a natural frequency ω_i and a fraction of critical damping β_i for the combination of the isolator and the mass. By definition:

$$\omega_i = \sqrt{k_i/m}, \quad \beta_i = 2\sqrt{k_i m} \quad (10)$$

The real and imaginary parts of the isolator ad-

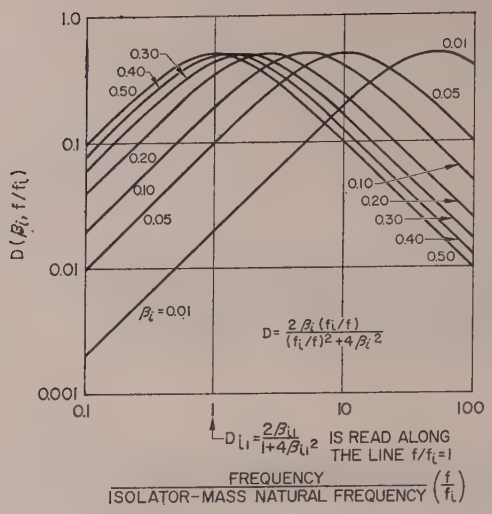


Fig. 2 — Plot of D function versus f/f_i for constant β_i (Step B7 of Design Procedure)

[5] As used here, the terms "admittance" and "mobility" are synonymous.

[6] As the source admittance approaches zero, or the source impedance approaches infinity, H approaches the commonly defined transmissibility. See footnote 7.

[7] "Vibration and Shock Control," by C. E. Crede, p. 37. Pub. by John Wiley and Sons, Inc., New York, 1951.

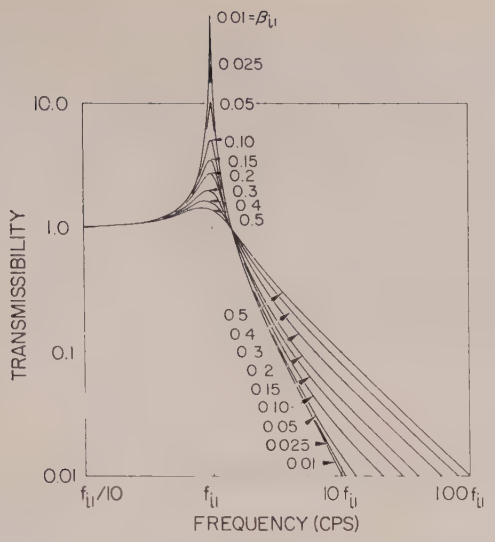


Fig. 3 — Family of classical transmissibility curves, which serve as approximation of H_i transmissibility in neighborhood of its first peak (Steps B2 and B8(c) of Design Procedure)

mittance may be rewritten in terms of these parameters with the result:

$$A_i(\omega) = \frac{1}{m\omega_i}\frac{\omega}{\omega_i} D(\omega/\omega_i, \beta_i) \tag{11}$$

$$B_i(\omega) = \frac{1}{m\omega_i}\frac{1}{2\beta_i} D(\omega/\omega_i, \beta_i) \tag{12}$$

where:

$$D(\omega/\omega_i, \beta_i) = \frac{2\beta_i(\omega_i/\omega)}{(\omega_i/\omega)^2 + 4\beta_i^2} \tag{13}$$

is a dimensionless parameter which can be provided in the procedure by a family of curves of $D(\omega/\omega_i, \beta_i)$ versus ω/ω_i for a range of values of β_i (Fig. 2).

Another H transmissibility to be defined is derived from the equipment fragility curve $V_f(\omega)$, which is a plot of the amplitude of sinusoidal velocity which the equipment can withstand versus frequency. By implication, if the equipment velocity exceeds the fragility velocity at any frequency, an uncertainty exists as to whether the equipment will fail and, if it does, the condition must be termed unsatisfactory. There results a maximum allowable transmissibility.

$$H_m(\omega) = \left| \frac{V_f(\omega)}{V_s(\omega)} \right| \tag{14}$$

The criterion for a satisfactory installation is that $|V(\omega)| \leq |V_f(\omega)|$ at all frequencies, which is equivalent to the criterion:

$$H(\omega) \leq H_m(\omega) \text{ for all } \omega \rightarrow \begin{matrix}\text{satisfactory} \\ \text{installation}\end{matrix} \tag{15}$$

On the other hand, if $|V(\omega)| > |V_f(\omega)|$ at any ω, the installation is unsatisfactory, yielding the criterion:

$$H(\omega) > H_m(\omega) \text{ for any } \omega \rightarrow \begin{matrix}\text{unsatisfactory} \\ \text{installation}\end{matrix} \tag{16}$$

Synthesis of Generalized Transmissibility — The initial choice of isolators in the design procedure is based on the maximum allowable transmissibility at frequencies below the first resonant frequency of the unloaded source.[8] Behavior at higher frequencies is ignored until later in the procedure. It is first assumed that the supporting structure and isolator combination behaves as a parallel spring-dashpot element in this low frequency range. The system composed of the equipment mass supported on this fictitious spring-dashpot element is then a single degree of freedom system for which the classical transmissibility curves in Fig. 3 apply. Various combinations of the undamped natural frequency and fraction of critical damping for this simple system can then be chosen so that the maximum allowable transmissibility is not exceeded. This process is carried out graphically, by using Fig. 3 as an overlay on H_m, so that a region of natural frequency and damping combinations is readily produced.

The problem then is to find the isolator parameters required, in conjunction with the admittance of the source, to yield the desired behavior of the overall system. The problem is one of synthesis, since it is desired to find the properties of a portion of the system in order that a given characteristic is achieved for the overall system.

If it is assumed that in the neighborhood of the lowest resonant frequency the system behaves precisely as a single degree of freedom structure with mass m, undamped natural frequency ω_{i1} and fraction of critical damping β_{i1}, then the modified source admittance (Fig. 1C) at the frequency ω_{i1} is given by:

$$A_s(\omega_{i1}) + A_i(\omega_{i1}) = \frac{1}{m\omega_{i1}}\frac{2\beta_i}{1+4\beta_{i1}^2} \tag{17}$$

$$B_s(\omega_{i1}) + B_i(\omega_{i1}) = \frac{1}{m\omega_{i1}}\frac{1}{1+4\beta_{i1}^2} \tag{18}$$

The right sides of Eqs. 17 and 18 are exactly the real and the imaginary parts, respectively, of the admittance of a spring-dashpot combination supporting a mass m, which would yield an undamped natural frequency ω_{i1} and a fraction of critical damping β_{i1}. Thus, knowing the first mode resonance parameters and the source admittance at ω_{i1}, the isolator admittance at ω_{i1} can be found readily by use of Eqs. 17 and 18. Then Eqs. 11–13 can be solved for ω_i, and β_i when $\omega = \omega_{i1}$, with the result that:

$$\omega_i = \frac{\omega_{i1}}{m}\sqrt{\frac{B_i(\omega_{i1})}{A_i^2(\omega_{i1}) + B_i^2(\omega_{i1})}} \tag{19}$$

and:

$$\beta = \frac{1}{2}\frac{\omega_i A_i(\omega_{i1})}{\omega_{i1} B_i(\omega_{i1})} \tag{20}$$

Consequently, the required isolator parameters ω_{i1}, β_i are obtained from the system parameters ω_{i1}, β_{i1}.

Some of the ω_{i1},β_{i1} combinations which provide a satisfactory transmissibility characteristic may not be obtainable with spring-dashpot isolators because of the limitation that neither the real or the imaginary parts of the isolator admittance can be negative. This limitation can be applied in the solution of Eqs. 17 and 18 to yield a maximum possible β_{i1} by setting $B_i(\omega_{i1}) = 0$ in Eq. 18.

$$\beta_{i1}(\max) = \frac{1}{2}\sqrt{\frac{1}{m\omega_{i1}B_s(\omega_{i1})} - 1} \tag{21}$$

and to yield a minimum possible β_{i1} by setting

[8] The lowest resonant frequency of the unloaded source is located at the first zero in the imaginary part of the source admittance.

$A_i(\omega_{i1}) = 0$ in Eq. 17:

$$D_{i1} \text{ at } \beta_{i1}(\min) = m\omega_{i1}A_s(\omega_{i1}) \tag{22}$$

where:

$$D_{i1} = \frac{2\beta_{i1}}{1 + 4\beta_{i1}^2} \tag{23}$$

Upper Bound to Transmissibility at High Frequencies — Once the transmissibility at low frequencies has been controlled, it is desirable to investigate the possibility of exceeding the maximum allowable transmissibility at high frequencies. To determine quickly whether a possibility of trouble exists, an upper bound to the high frequency transmissibility is found. This upper bound transmissibility, $H_i(\omega)$, is given by:

$$\overline{H}_i(\omega) = \frac{1}{m\omega[A_s(\omega) + A_i(\omega)]} \tag{24}$$

which is exactly the transmissibility when the term $\{1 - m\omega[B_s(\omega) + B_i(\omega)]\}$ in Eq. 6 equals zero.[9] If the term is nonzero, either in a positive or a negative sense, the transmissibility is smaller. An alternate expression utilizing Eq. 11:

$$\overline{H}_i(\omega) = \frac{1}{\frac{\omega^2}{\omega_i^2} D(\omega/\omega_i, \beta_i) + m\omega A_s(\omega)} \tag{25}$$

yields values more easily, since a plot of $D(\omega/\omega_i, \beta_i)$ is available.

If it is desired to check an isolator design quickly for possible trouble at high frequencies, values of $\overline{H}_i(\omega)$ for that design can be calculated over the entire high frequency range. If $\overline{H}_i(\omega) < H_m(\omega)$, there is no possibility of trouble at high frequencies; if $\overline{H}_i(\omega) > H_m(\omega)$ at any frequency, the isolator design is suspect at that frequency. Suspect designs can be examined further by determining whether or not any high frequency peaks are present in the regions within which $\overline{H}_i(\omega) > H_m(\omega)$. If no peaks exist in the suspect regions, $H_m(\omega)$ probably is not exceeded. However, this can not be stated with certainty since $H_m(\omega)$ can be exceeded at frequencies other than those at which peaks occur, although this would be unusual. Often, enough of a range of nonsuspect designs exists so that there is no interest in further investigation of the suspect designs.

Additional Expressions — In order not to waste effort on isolators so flexible that collisions of the equipment with neighboring structures or collisions internal to the isolators will occur, a lower bound is placed in the isolator stiffness, $k_i(\min)$. This lower bound can be determined by requiring that low frequency accelerations of the supporting structure not cause excessive motion of the equipment relative to its supports.

If $a_{\max}$ is the largest peak acceleration[10] of the supporting structure (obtained from V_s) in in./sec^2 in an arbitrarily established low frequency region, say < 10 cps, and $u_{\max}$ is the maximum permitted displacement across the isolators in inches measured from the static equilibrium position of the equipment, the minimum stiffness of each individual isolator is:

$$k_i(\min) = \frac{ma_{\max}}{nu_{\max}} \tag{26}$$

where:

n = Number of identical isolators used

Eq. 26 assumes that no dynamic magnification occurs, and is useful because it can be applied before the designs are chosen from the transmissibility requirements. After the designs are chosen, the degree to which the isolator magnifies the motion of the supporting structure should be examined to determine if a further restriction on isolator stiffness exists due to equipment displacement limitations.

In the procedure, the stiffness of the isolator is expressed in terms of a natural frequency ω_i in association with the equipment mass m. For use in the procedure, the relation:

$$\omega_i(\min) = \sqrt{\frac{a_{\max}}{u_{\max}}} \tag{27}$$

is a limitation entirely equivalent to Eq. 26.

After isolator designs expressed in terms of ω_i and β_i are chosen, it may be desirable to find the stiffness and viscous damping coefficient corresponding to ω_i, β_i for each individual isolator. This can be accomplished with the relations:

$$nk_i = m\omega_i^2 \tag{28}$$

and:

$$nc_i = 2m\beta_i\omega_i \tag{29}$$

These relations are obtained from the basic definitions of natural frequency and fraction of critical damping given in Eq. 10.

Data Required

In order to apply the design procedure in a given direction of vibrational motion, the data given below for the given direction are required. The units are force in pounds, length in inches, and time in seconds.

It is clear that whenever the ratio of two motions is involved, it is necessary only that both motions be of the same type and in the same units. Amplitudes of sinusoidal motion may be expressed in terms of displacement, velocity, or accelerations. Velocity amplitudes are used in order to be definite in discussing the procedure. All frequencies will be expressed in cycles per second (cps) so that whenever ω appears in a relation, $2\pi f$ will replace it in the procedure.

1. Weight of the equipment in pounds W, where $W = mg$

2. Fragility of the equipment, $V_f(\omega)$

The fragility of the equipment $V_f(\omega)$ is expressed by a plot of sinusoidal amplitude versus frequency, determined in such a manner that failures are highly unlikely if sinusoidal motions of lower amplitudes are applied to the equipment. Such a plot can be derived by sinusoidal vibration testing of the equipment. It is assumed that the fragility plot has been determined in a manner which considers the difference in time between the test and the environmental exposure to the vibration, the degree of confidence desired in the application of the plot, and the increased likelihood of failure due to superposition effects among the sinusoidal components. In some cases it may be necessary to supplement the low frequency portion of the fragility curve with centrifuge

[9] The term $\{1 - m\omega[B_s(\omega) + B_i(\omega)]\}$ goes to zero at the undamped resonant frequencies of the system. Sharp high frequency peaks, therefore, will rise to touch the $\overline{H}_i(\omega)$ curve.

[10] Exclusive of the gravitational acceleration.

testing and/or calculated behavior below the first resonant frequency of the equipment.

3. Free (unloaded) vibration of source, $V_s(\omega)$

The free vibration of the source, $V_s(\omega)$, is determined by measuring the vibrational motion existing on the structure to which the equipment is to be attached when the equipment is not mounted.[11] The relatively light vibration pickup used to make the measurement usually has an insignificant effect on the no-load source motion. At each frequency it is assumed that the amplitudes of vibration of the various attachment points are enough alike so that motion of an installed equipment will be principally translational. The amplitude of the single equivalent source, as shown in Fig. 1, is then taken to be the average of the amplitudes at the attachment points. A somewhat more refined method of deriving the equivalent amplitude is given in the following section on the admittance of the source.

Under any single set of operating conditions, it is assumed that the vibration has a line spectrum. Hence, the motion $V_s(\omega)$ occurs at a discrete set of frequencies. Ordinarily, however, the equipment must operate under a variety of conditions. If the discrete sets of sinusoidal vibrations associated with all the various conditions are considered, the result is a series of discrete frequency bands, each of which shows the variation of one of the discrete sinusoidal components with changes in operating conditions. Where such bands overlap, the distribution of sinusoidal components becomes continuous.

4. Mechanical admittance of source, $Y_s(\omega) = A_s(\omega) + jB_s(\omega)$

If a sinusoidal force is applied to an attachment point of the supporting structure, a sinusoidal velocity at the same frequency will result. With no load present on the source, the ratio of the complex amplitude of this resulting velocity to the complex amplitude of the driving force is termed the "source mechanical admittance" at that frequency. Two curves are required for the procedure, one of the real part $A_s(\omega)$, and the other the imaginary part $B_s(\omega)$.

More often the mechanical impedance $Z_s(\omega) = R_s(\omega) + jX_s(\omega)$, the reciprocal of admittance, is considered. However, for the purposes of design procedure, it has been found advantageous to use the mechanical admittance because the isolators are series connected to the source, and thus produce a modified source whose admittance is the sum of the admittance of the source and of the isolator.

If at each frequency the resultant of the forces delivered to the equipment by the motions of the supporting structure passes through the equipment center of gravity, pure translation motion results. In order for this to be valid over a large frequency range, both the free motions and the admittances at the various equipment support points must be equal in that frequency range, the support points must be symmetrically located with respect to the equipment center of gravity and, if isolators are used at each support, they must be identical. These conditions apply separately in each direction of motion.[12]

Assuming that these conditions are satisfactorily met, the impedance of the equivalent single source (Fig. 1) is the sum of the impedances at the various mounting points, namely:

$$Z_s(\omega) = \sum_n Z_{sn}(\omega) \qquad (31)$$

where:

$Z_{sn}(\omega)$ = Impedance at n^{th} mounting point.

Impedances are cumulative since the mounting points act in parallel; in other words, a force applied to the equipment is resisted by the sum of the forces at the various mounting points. The source admittance required for the procedure is thus:

$$Y_s(\omega) = 1/\sum_n Z_{sn}(\omega) = 1/\sum_n [1/Y_{sn}(\omega)] \qquad (32)$$

where:

$Y_{sn}(\omega)$ = Admittance at n^{th} mounting point.

Eq. 32 is applied, even though the admittances at the mounting points are only approximately equal.

It was mentioned previously that a more refined method than merely averaging the approximately equal free amplitudes at the various mounting points is available for finding the single equivalent free source amplitude. It is even better to require that the absolute value of the dynamic force delivered by the single equivalent source be the sum of the absolute values of the forces delivered by the various mounting points. Referring to Eq. 1, the sum of absolute values of the forces delivered by the n mounting points is:

$$\sum_n |F_n(\omega)| = \sum_n |Z_{sn}(\omega)[V_{sn}(\omega) - V_n(\omega)]|$$

If the amplitude of the force delivered by the single equivalent source, $F(\omega)$, is made equal to $\sum_n |F_n(\omega)|$, and the performance of the single source is as given by Eq. 1, then it can be shown that:

$$|V_s(\omega)| = \frac{\sum_n |Z_{sn}(\omega) V_{sn}(\omega)|}{\sum_n |Z_{sn}(\omega)|} \qquad (33)$$

It can be seen that if the $Z_{sn}(\omega)$ are equal, $|V_s(\omega)|$ is the average of the $|V_{sn}(\omega)|$. Since, if the assumption of no significant rotation of the equipment is

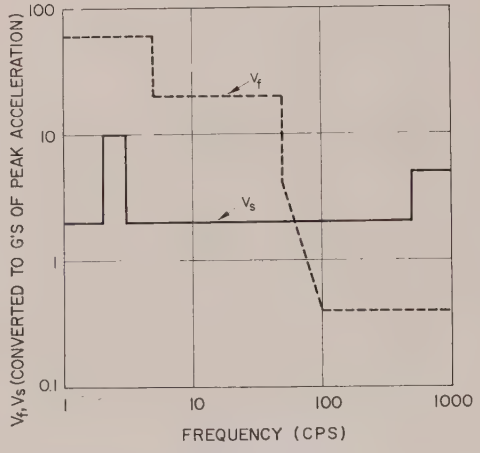


Fig. 4 — Plots of fragility amplitude V_f and free source amplitude V_s (items 2 and 3 of Data Required)

[11] If for some reason a no-load condition is impractical, the measruement still can be made with a known load present. Referring to Fig. 1A, if the input admittance of the load is $Y_L(\omega)$, and the source output vibration measurement at a frequency ω is $V(\omega)$, then the corresponding source free vibration is given by an equation derived from Eq. 1 and the condition that $Y_L(\omega)F(\omega) = V(\omega)$.

$$V_s(\omega) = \left[1 + \frac{Y_L(\omega)}{Y_s(\omega)}\right] V(\omega) \qquad (30)$$

[12] This situation is sufficient to produce pure translation, but is not required. Pure translation also can be produced by nonsymmetrical locations of the support points if the free motions at the support points are equal and if the sum of the moments of the modified admittances (per Eq. 5), with respect to each axis through the center of gravity, is zero.

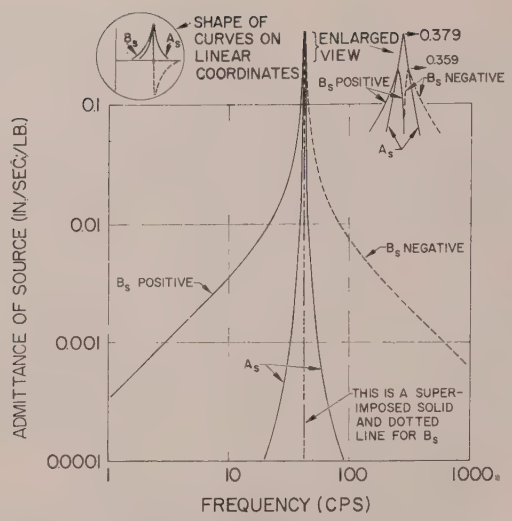


Fig. 5 — Mechanical admittance of source (item 4 of Data Required and Step A1 of Design Procedure)

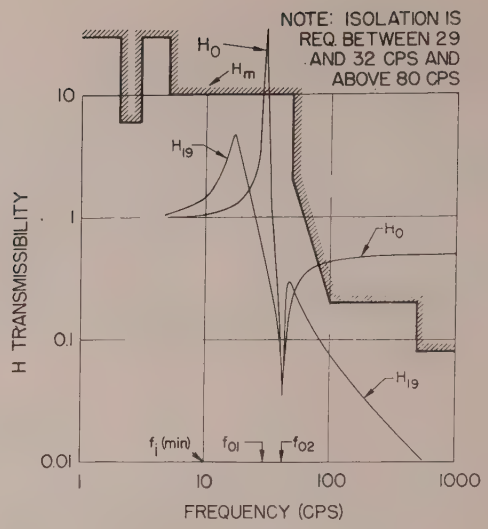


Fig. 6 — Plots of H transmissibility (Steps A2-A4 and B1 and B7 of Design Procedure)

valid, the $Z_{sn}(\omega)$ can not be too different, the average free motion is probably a satisfactory approximation. If any doubt exists, Eq. 33 should be used.

5. Maximum low frequency acceleration $a_{\max}$, and maximum allowable displacement u_{max}

The maximum low frequency acceleration of the supporting structure and the maximum allowable displacement across the isolators are required only after it has been determined that isolation is needed. (Refer to Eq. 26 for the meaning of these parameters.)

Design Procedure

The steps involved in the design procedure as it now exists are outlined here. The procedure should be applied in each of three mutually perpendicular directions taken one at a time. The minor procedural details are omitted for brevity, and only the major steps are stated. The solution of a sample problem, defined by the following data, is included.

Data: $W = 100$ lb
$V_f(\omega)$ and $V_s(\omega)$ are as given in Fig. 4[13]
$A_s(\omega)$ and $B_s(\omega)$ are as given in Fig. 5
$a_{\max} = 10g = 3860$ in./sec^2
$u_{\max} = 1$ in.

A. Determining Need for Vibration Isolation — The first stage of the procedure is to determine whether or not vibration isolation is required. This is done by assuming that the equipment is mounted directly (without isolators) to the supporting structure. No isolation is required if the amplitude of the equipment motion is less than the fragility amplitude at all frequencies. The amplitude of the equipment motion in terms of the generalized transmissibility function H is calculated in steps, as follows.

1. Plot the real and imaginary parts of the source admittance, $A_s(\omega)$ and $B_s(\omega)$, on log-log graph paper, using solid curves for positive values and dotted curves for negative values of $B_s(\omega)$. Negative values of $A_s(\omega)$ can not occur. The result for the sample problem is shown in Fig. 5.

2. Calculate and plot the ratio of the fragility amplitude, $V_f(\omega)$, to the free source amplitude, $V_s(\omega)$, and call this $H_m(\omega)$, the maximum permitted transmissibility. The result from the data in Fig. 4 appears in Fig. 6 with cross-hatching implying that it is an upper bound.

3. Calculate $H_0(\omega)$, the transmissibility when the equipment is mounted directly to the source, using Eq. 3′.[14] The family of curves shown in Fig. 7 can be used to facilitate the solution of Eq. 3′. By calculating $m\omega A_s(\omega)$ and $|1 - m\omega B_s(\omega)|$, $H_0(\omega)$ can be read directly from Fig. 7 when a scale factor is properly applied.[15] The lowest frequency peak is labeled ω_{01}, the next valley ω_{02}, the next peak ω_{03}, and so on.

4. Plot $H_0(\omega)$ on the same graph as $H_m(\omega)$. (See Fig. 6) Apply the criteria of Eqs. 15 and 16. That is, if $H_o > H_m$ at any frequency isolation is required and the procedure continues; if $H_0 < H_m$ at all frequencies, no isolation is required and the procedure stops. It is seen that for the sample problem isolation is required.

B. Determining Required Isolator Parameters — If isolation is required, the following steps lead to the determination of those isolator parameters which will provide a satisfactory installation:

1. Using Eq. 27, determine a minimum natural frequency, $\omega_i(\min)$, for the isolation system. Locate this frequency on the frequency axis of the graph on which H_m and H_0 have been plotted. (See Fig. 6 on which the value $\omega_i\,(\min) = \sqrt{3860/1} = 62.2$ rad/sec = 9.9 cps has been plotted.)

2. Determine the region of permissible damping and natural frequency parameters, β_{i1} and ω_{i1}, for the lowest frequency peak in the $H_i(\omega)$ transmissibility curves defines by Eq. 6′. As an approximation to $H_i(\omega)$ in the neighborhood of ω_{i1}, use the family of curves shown in Fig. 3.[16] With the re-

[13] The ordinates in Fig. 4 are g's of peak acceleration. It was done this way to emphasize the point that it is only required that $V_f(\omega)$ and $V_s(\omega)$ be expressed in the same terms. Only their ratio appears in the procedure.

[14] In the present version of the procedure a method has been provided to determine the minimum number of frequencies at which $H_0(\omega)$ should be calculated in order that it is properly defined. The method graphically locates the peaks and valleys of $H_0(\omega)$ and then determines additional intermediate frequencies at which calculations should be made.

[15] The multiplication of both $m\omega A$ and $|1 - m\omega B|$ by 10^n results in a multiplication of H by 10^{-n}.

[16] The curves in Fig. 3 are the classical transmissibility curves for a single degree of freedom isolator-mass configuration. A transparency of this family of curves will be available for the use as an overlay on the plot of $H_0(\omega)$ and $H_m(\omega)$.

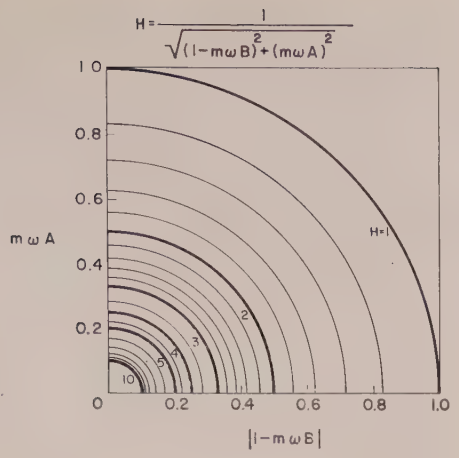


Fig. 7 — Curves of $m\omega A$ versus $|1 - m\omega B|$ for constant H, used to facilitate solution of Eqs. 3′ and 5′ (Step A3 of Design Procedure)

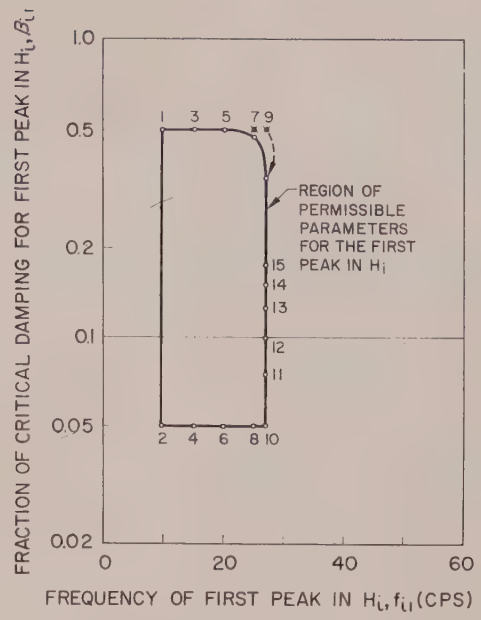


Fig. 8 — Parameters for first peak in H_i curve (Steps B2-B3 of Design Procedure)

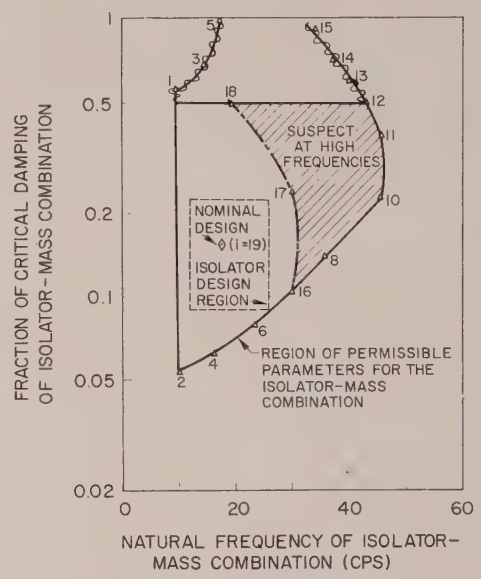


Fig. 9 — Parameters for isolator-mass combination (Steps B4-B6 of Design Procedure)

striction that $\omega_i(\min) \leq \omega_{i1} < \omega_{01}$, and ignoring the frequency range above ω_{02}, establish a region of ω_{i1}, β_{i1}, combinations for which the transmissibility in Fig. 3 lies below H_m.[17] Label this "Region of Permissible Parameters for the First Peak." For the sample problem, see Fig. 8 for the region which results. Various designs on the boundary of the region are numbered for future reference.

3. Eliminate that portion of the "Region of Permissible Parameters for the First Peak" which can not be achieved with spring-dashpot isolators. Use Eqs. 21 and 22 to find maximum and minimum values of β_{i1} for the various ω_{i1}. Remove those designs which lies outside these limiting values. In the sample problem the values of β_{i1} for the two designs at the upper right-hand corner of the region in Fig. 8 lie above $\beta_{i1}(\max)$. These two designs are removed and designs immediately below them at the values of $\beta_{i1}(\max)$ are used to form a changed corner of the region.

4. Determine the region of permissible damping and natural frequency parameters for the isolators, β_i and ω_i. That is, determine the region of β_i, ω_i which corresponds to the region of permissible β_{i1}, ω_{i1} which was established in the paragraphs above, by applying Eqs. 17–20 to designs on the boundary of the β_{i1}, ω_{i1} region. Eliminate those designs for which $\beta_i > 0.5$ as being impractical.[18] The resulting region for the sample problem, called the "Region of Permissible Parameters for the Isolators," is shown in Fig. 9. The numbered designs on the boundary of the region correspond to the like-numbered designs on the boundary of the region in Fig. 8. (Designs 16, 17, 18 are added for future reference.)

5. Find those permissible isolator designs which are suspect at high frequencies. The permitted designs found in the previous paragraph were determined without considering the frequency region above ω_{02}, the location of the first minimum in the $H_0(\omega)$ curve and also the location of the first resonance of the unloaded source. This high frequency region will now be investigated. In conjunction with a graphical procedure, Eq. 25 can be used to define the limits of the suspect designs. Details are omitted for brevity. The suspect designs are those for which $\overline{H}_i(\omega)$ exceeds $H_m(\omega)$ at some frequency. The suspect region for the sample problem is shown in Fig. 9.

6. Choose a nominal isolator design within the region of permissible designs which meets the following conditions:

(a) A reasonable variation in natural fre-

[17] When B_s is spring-like at low frequencies, namely, when B_s is positive and proportional to frequency, H at $\omega = 0$ will equal one and the above procedure is satisfactory. When B_s is mass-like at low frequencies, namely, when B_s is negative and inversely proportional to frequency, H at $\omega = 0$ will equal $1/(1 + m/m_s)$. The quantity m_s is the effective mass of the source at low frequencies, and m is the mass of the equipment. In this event, the above procedure may have to be modified. If the equipment being considered is airborne, m_s is the mass of the entire carrying vehicle. The ratio of equipment to vehicle mass is usually quite small, with the result that the effect of m_s on H is insignificant at low frequencies, and the above procedure is satisfactory. If $m/m_s > 0.1$, the procedure will probably require modification.

[18] This limitation is applied at this stage of the procedure so that the family of curves in Fig. 2 can be used to facilitate further calculations. In the region of $\beta_i > 0.5$ the curves in Fig. 2 do not exist.

quency ω_i, for example, ± 10%, can be tolerated.

(b) A reasonable variation in fraction of critical damping β_i, for example, ± 20%, can be tolerated.

(c) Commercially available isolators can be used, or if not available, the design should be within the capabilities of the current state of the art in the production of vibration isolators.

(d) The design, including the tolerances in (a) and (c), attempts to avoid the suspect region. If the nonsuspect region is too limited, further investigation of the suspect designs is warranted. The calculation of $H_i(\omega)$ at the high frequency peaks is often sufficient to verify the suspicion or to remove it from a design.

A nominal design, designated by $i = 19$, is shown on Fig. 9 for the sample problem.

7. Calculate the transmissibility achieved using the nominal isolator design as a check to verify that $H_i(\omega) < H_m(\omega)$ at all frequencies. Calculate the isolator admittance as a function of frequency, using Eq. 11 and 12 and Fig. 2.[19] Add the isolator and source admittance to obtain a modified source admittance. The procedure used in Step 3 to obtain $H_0(\omega)$ can now be used to plot $H_i(\omega)$ on the graph containing $H_m(\omega)$ and $H_0(\omega)$. The resulting $H_{19}(\omega)$ for the sample problem is shown in Fig. 6.

8. State the isolator requirements by one or more of the three methods below. The region of permissible isolator parameters forms the basis for the statement of isolator requirements. If the designs to be used are suspect, they must be checked to insure that their high frequency behavior is satisfactory.

(a) On the basis of a nominal design having the attributes stated in Step 6, state the requirements as follows:
n isolators are required such that the isolator-equipment combination has a resonant frequency of 19 ± 7 cps and a fraction of critical damping of 0.15 ± 0.06. The results obtained apply to the sample problem; the design region is shown in Fig. 9. The numbers are independent of n.

(b) Restate the requirements in (a) in terms of isolator stiffness and viscous damping coefficient as follows:
Four isolators are required such that individually their equivalent dynamic stiffness is 1050 ± 680 lb/in. and their individual coefficient of viscous damping is 2.67 ± 1.78 lb/in./sec. Eqs. 28 and 29 are used to make the transformation. The results obtained apply to the sample problem on the basis that $n = 4$.

(c) Establish an envelope of permissible transmissibility curves for the isolator-equipment combination.[20] This can be done most easily by tracing a series of curves from Fig. 3 on a log-log graph. Each curve results from a design on the boundary of permissible isolator designs. The individual curves are numbered to correspond to numbered designs in Fig. 9. The natural frequency ω_{i1} on Fig. 3 corresponds to the isolator equipment natural frequency, ω_i. The maximum allowable transmissibility for the sample problem is shown in Fig. 10.

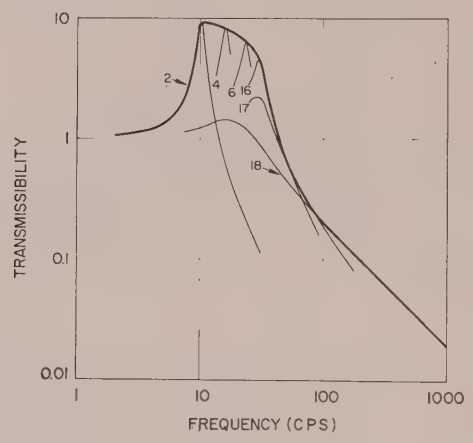


Fig. 10 — Maximum allowable transmissibility for isolator-mass combination (Step B8 (c) of Design Procedure)

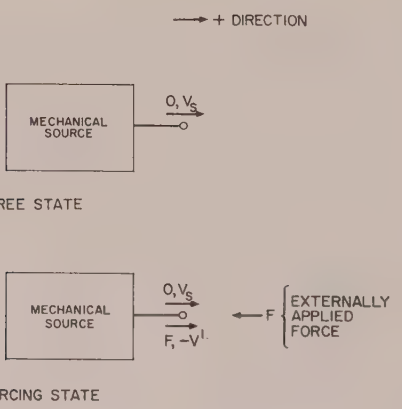


Fig. 11 — States of simple mechanical source

Conclusion

A procedure is presented for the dynamical design of center-of-gravity vibration isolation suspensions for fragile equipment when supported by a nonrigid structure which is subject to sinusoidal vibrations. At its present stage of development the procedure has been used to solve some 10 problems covering a variety of parameters. It is concluded that the mechanical admittance, or mobility of the supporting structure, is often a highly significant parameter. If admittance data is not available because it has not been measured or estimated analytically, the procedure still can be applied by assuming that the admittance is zero (infinitely rigid). With this simplification, the labor in applying the procedure is considerably reduced, since the classical transmissibility curves in Fig. 3 apply at all frequencies.

Although the procedure is designed to be carried out with the use of a slide rule and/or desk calculator, there is no reason why the procedure cannot be modified for use with high-speed digital computing equipment.

Some minor modifications of the procedure are

[19] A graphical method has been provided in the present version of the procedure to locate the peaks and valleys of $H_i(\omega)$, which are then used to determine additional intermediate frequencies at which calculations should be made in order to minimize the number of points needed to define $H_i(\omega)$.

[20] This envelope is the maximum allowable transmissibility when the source impedance is considered to be infinite. It applies to transmissibility measurements taken when the isolator-equipment combination is tested on a vibrator. This type of statement of the requirements has the advantage that it can be applied directly to slightly nonlinear isolators, obviating the determination of linearized isolator parameters.

still required, mainly in the area of simplifying and clarifying the details for ease in use. Although further applications of the procedure may reveal deficiencies, it is felt that these can be overcome within the existing framework. The SAE G-5 committee is presently attempting to extend the procedure to include random input motions of the supporting structure, after which shock motions probably will be considered. The committee is striving to treat both of these types of inputs in a manner that will entail minimum alteration of the present method of attack.

Acknowledgment

The author wishes to express his appreciation to Dr. C. T. Molloy, who provided the philosophical framework for this work; the many discussions with him were most helpful. Thanks are also expressed to H. Himelblau, who presented his class with a number of test problems to be solved by the procedure, and to the members of his class who devoted a great deal of time to the solutions. In addition, the efforts of those members of the G-5 committee who attended work sessions and who provided constructive criticism are gratefully acknowledged.

APPENDIX

Performance of Simple[21] Mechanical Source

In Molloy[4] the performance equation of a simple source is derived by the use of four-pole parameters. A completely equivalent derivation may be achieved merely by the use of superposition.

The following discussion applies to each frequency at which sinusoidal motion exists. A beginning may be made with a source in its free state, as shown in Fig. 11A, where the output velocity is V_s and the output force is zero. The application by some external agent of a force, F, to the output terminal in the negative direction, as shown in Fig. 11B, is equivalent to the condition that the source is applying a positive force F to the load; this may be termed a forcing state of the source. The velocity of the output terminal in response to the external force F is in a negative direction, and its amplitude is V'. In terms of the applied force:

$$V' = F/Z_s \tag{34}$$

where:

Z_s = Point impedance of source at its output terminal

Assuming that V_s is solely determined by the active elements within the source, that is V_s is not altered by the presence of F, then the total velocity of the output terminal, as shown in Fig. 1A, is $V = V_s - V'$. Substituting the definition of V' given in Eq. 34 and rearranging the terms:

$$F = Z_s(V_s - V) \tag{1}$$

which is the desired result.

Nomenclature

A = Real part of mechanical admittance, in./sec./lb
B = Imaginary part of mechanical admittance, in./sec/lb
D = Dimensionless function defined in Eq. 13 and plotted in Fig. 2; used to facilitate calculation of admittance of isolator
D_{i1} = Dimensionless function defined in Eq. 23; special case of D
F = Complex amplitude of sinusoidal force, lb
H = Generalized transmissibility function defined by Eqs. 7 and 8 (dimensionless)
$\bar{H}$ = Upper bound to H – defined by Eq. 24
R = Real part of mechanical impedance, lb/in./sec
V = Complex amplitude of sinusoidal velocity, in./sec
W = Weight of equipment, $W = mg$, lb
X = Imaginary part of mechanical impedance, lb/in./sec
Y = Complex mechanical admittance, in./sec/lb
Z = Complex mechanical impedance, lb/in./sec
a_{max} = Largest peak acceleration of supporting structure at very low frequencies, in./sec^2
c = Viscous damping coefficient, lb/in./sec
f = Frequency, $f = \omega/2\pi$, cps
g = Acceleration of gravity, 386 in./sec^2
$j = \sqrt{-1}$
k = Linear spring stiffness, lb/in.
m = Mass of equipment, lb/in./sec^2
n = Number of support points or number of isolators
u_{max} = Maximum permitted displacement across isolators, in.
β = Fraction of critical damping (dimensionless)
ω = Frequency, $\omega = 2\pi f$, rad/sec

Subscripts

0 — Subscript zero used with F, H, and V; indicates that equipment is attached to source without use of isolators
01, 02 — Used with ω and f to indicate frequencies at which first peak 01, and first valley 02 occur in H_0
f — Used with V to indicate so-called fragility amplitude of equipment
i — Used with c, k, A, B, and Y to indicate that quantities refer to isolators
Used with β, ω, and f to indicate fraction of critical damping and undamped natural frequency of isolator-mass combination
Used with H to denote that isolators are present between source and mass
$i1$ — Used with β, ω, and f to indicate fraction of critical damping and undamped natural frequency of first peak in H_i
Used with D to denote relation in Eq. 23
L — Used with Y to denote admittance of load on source
m — Used with H to denote maximum allowable transmissibility defined by Eq. 14
n — Used with F and V to denote quantities at n^{th} support point
s — Used with A, B, Y, and R, X, Z to denote admittance and impedance quantities of source structure
Used with V to denote velocity of source structure when no load is present (free velocity of source). A single equivalent source is referred to in all cases
sn — Used with same quantities that use subscript s; denotes quantities at n^{th} support point

[21] The word "simple" implies that only a single output terminal is involved. A more complicated source involves a multiplicity of output terminals which do not behave independently.

SPARK-PLUG MISFIRING —mechanism studies

H. P. Julien and R. F. Neblett
Esso Research and Engineering Co.

FOR MANY YEARS spark-plug misfiring has been of considerable interest to both the automotive and petroleum manufacturers. Many publications have appeared from both these sources discussing the extent of misfiring, how it occurs, and how it can be alleviated. (See Additional References.)

Spark-plug misfiring is generally caused by one of three conditions. In *gap bridging,* a carbonaceous deposit forms across the gap of the spark plug. This, of course, prevents firing of the plug at all speeds and under all operating conditions. Another condition which has been observed is *gap erosion.* The spark-plug electrodes are eroded by the electrical discharge and corrosive gases under the high-temperature conditions present in the combustion chamber; the gap widens to the point where the available voltage is not sufficient to fire a spark.

However, *lead fouling,* caused by the deposition of lead salts on the ceramic insulation of the plug, is the most common type of spark-plug fouling in passenger cars. Misfiring due to lead fouling has certain specific characteristics: (1) it occurs mainly at high speed, where the engine wall temperatures and the spark-plug ceramic temperatures are highest; (2) it occurs only after a period of relatively low-speed service; and (3) it is a transient phenomenon.

It seems probable that no one type of spark plug will provide a solution to all types of plug difficulties. Hot plugs have been found best suited to minimize lead fouling and gap bridging, while cold ones are best to reduce gap erosion (and surface ignition). On the other hand, lead fouling and gap bridging are problems which would seem to be susceptible to reduction by means of changes in fuel composition.

Mechanism Studies

In the mechanism of misfiring due to lead fouling as postulated by Lamb,[1] spark-plug temperature plays an extremely important role. During deposit accumulation, spark-plug temperature and the fuel components control the composition of the deposits. When more severe conditions of speed and/or load are imposed (as in an acceleration to high speed), spark-plug temperature and deposit composition determine the shunt resistance of the deposits. If this shunt resistance falls to some value critical to the engine and plug combination, misfiring occurs.

The shunt resistance below which a spark plug will misfire depends on many factors, including available voltage, gap size, gas pressure, and the concentration and nature of the gases between the plug electrodes. The relationship between low shunt resistance and misfiring is illustrated by the data obtained in a laboratory V-8 engine (Table 1). Shunt resistance was measured with a unit which measures the shunt resistance of a spark plug under actual operating conditions. Misfiring during a part throttle acceleration was observed on a DuMont engine analyzer. Plug resistance was determined from peak height on an oscilloscope.

Shunt resistance also depends upon the chemical composition of the deposits found on the spark plug, since the reduction of electrical resistance with increasing temperature varies for different lead compounds.

Both deposit composition and spark-plug temperature vary from cylinder to cylinder. It is not surprising, then, that spark-plug misfiring varies among the cylinders. When misfiring occurs, the spark-plug deposits cool off and simultaneously their shunt resistance increases. When the resistance

[1] "Postulated Mechanism of Spark Plug Fouling," by F. W. Lamb. Paper presented at 1954 Aircraft Spark Plug and Ignition Conference, October, 1954.

SPARK-PLUG MISFIRING due to lead fouling occurs when compounds of low electrical resistance are deposited on the spark-plug ceramic under low-temperature operating conditions and are subsequently raised to high temperature by operating the engine at higher speeds and loads.

Spark-plug temperature (as measured with thermocouple spark plugs) was found to be the most important variable governing misfiring. Deposit composition together with spark-plug temperature governs the shunt resistance of the spark plug. When this shunt resistance becomes sufficiently low, misfiring occurs.

Misfiring severity was shown to decrease when:

1. Deposits were accumulated at successively higher temperatures.

2. Engine conditions during deposit buildup were varied to yield higher maximum spark-plug temperatures although the average temperature was unchanged.

3. A phosphorus compound was added to the deposit accumulation fuel.

4. Fuel octane quality was increased to eliminate knock during accelerations.*

rises above the critical value, the plug will fire again. Fig. 1 shows the misfiring performance of four of the eight cylinders of a laboratory V-8 engine at 80 mph, wide-open throttle. Even in one cylinder misfiring is a random event.

Thus, misfiring is an erratic phenomenon depending upon a plug's shunt resistance, which in turn is dictated by temperature and deposit composition.

Laboratory Studies

Lead Compound Formation on Spark Plugs

Analysis of Spark-Plug Deposits — Spark-plug deposits occur in visually distinct bands on the plug ceramic. These bands of compounds of different composition are a reflection of the large temperature gradient existing between the base and tip of the plugs. In our analytical work, samples of spark-plug deposits were taken at three to four discrete intervals along the ceramic before elemental and X-ray diffraction analysis. Ordinarily, a given compound was found in two successive bands, indicating that sampling had been done at sufficiently small intervals.

A compound must be present in a concentration of at least 5–10% in order to be identified by X-ray diffraction. The analysis is further complicated by the presence of three or more lead compounds on a single spark-plug ceramic insulator. Elemental analyses on composite samples were used to confirm compounds identified by X-ray diffraction and to obtain quantitative results. These two types of analyses were combined to determine the amount of each compound present. Typical results are summarized in Table 2.

Compounds Identified in Spark-Plug Deposits — The most common compounds identified in spark-plug deposits occurring in passenger-car operation are the lead halides and oxyhalides. These result from the tel and accompanying scavenger in conventional leaded fuels. Many complexes of lead

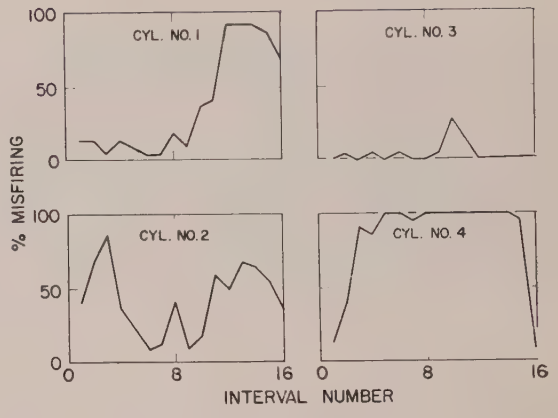


Fig. 1 — Misfiring pattern of individual cylinders of laboratory V-8 engine (per cent misfiring for 16 consecutive 1-sec intervals)

Table 1 — Misfiring Correlates with Shunt Resistance

Cylinder Number	Misfiring	Shunt Resistance, Megohms
2	Yes	0.8
5	Yes	1
3	Indefinite	3
8	Indefinite	3
7	Yes	5
1	No	> 5
4	No	> 10
6	No	> 10

Table 2 — Composition of Spark-Plug Deposit

	Lead Compounds, %		
	Tip	Middle	Base
$\alpha - 6PbO \cdot PbCl_2$	9	—	—
$2PbO \cdot PbClBr$	91	42	—
$PbClBr$	—	58	72
$PbCl_2$	—	—	28

* Paper presented at SAE National Fuels and Lubricants Meeting, Chicago, Oct. 28, 1959.

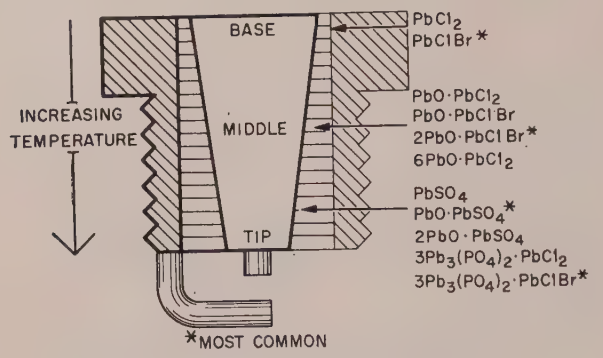


Fig. 2 — Many compounds have been identified in spark-plug deposits

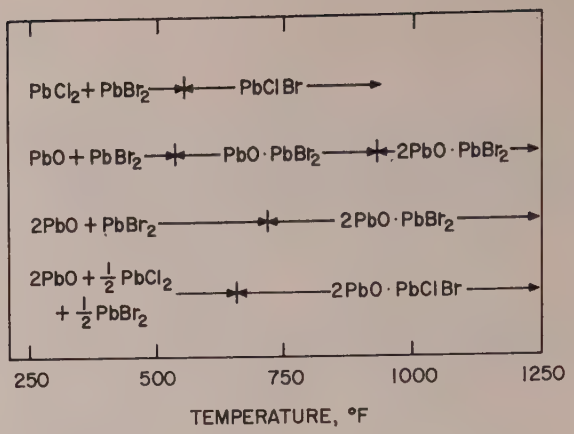


Fig. 3 — Temperatures of formation of lead oxyhalides

Table 3 — Reactions of Lead Compounds with P_2O_5 and H_3PO_4

Reactant	Reaction Products	
	With P_2O_5	With H_3PO_4
PbO	$5PbO \cdot Pb_3(PO)_4)_2$	$Pb_2P_2O_7$
	$PbO \cdot Pb_3(PO_4)_2$	$PbHPO_4$
	$Pb_2P_2O_7$	$Pb_3(PO_4)_2$
$PbCl_2$	$PbCl_2$	$PbCl_2$
	Pb	Pb
$PbBr_2$	$PbBr_2$	$PbBr_2$
	Pb	Pb
PbClBr	PbClBr	PbClBr
	Pb	Pb
$PbO \cdot PbClBr$	$3Pb_3(PO_4)_2 \cdot PbClBr$	$3Pb_3(PO_4)_2 \cdot PbClBr$
	PbClBr	PbClBr
		Pb
$PbO + PbCl_2$	$3Pb_3(PO_4)_2 \cdot PbCl_2$	$PbHPO_4$
		$PbCl_2$
		Pb

oxide with the lead halides have been found. Fig. 2 lists some of the compounds identified in deposits and indicates their normal location on the ceramic.

Lead sulfates and oxysulfates result from sulfur in the fuel. These compounds are found only on the tip and middle regions of the spark plug and they usually accompany high-temperature operation. When a phosphorus additive is incorporated in the fuel, a complex lead phosphate-lead halide, $3Pb_3(PO_4)_2 \cdot PbClBr$, is formed on all parts of the plug. Possible chemical reactions leading to the formation of the lead halide-lead phosphate complex were investigated in the laboratory. The reactants were heated in the open air in aluminum vessels at 800 F until visible reaction ceased. The products were identified by X-ray diffraction. The results are summarized in Table 3.

It is interesting to note that $3Pb_3(PO_4)_2 \cdot PbX_2$ (X = halogen), which is the same compound found in spark-plug deposits, is not formed from lead halide alone. The oxide must also be present, either combined or as part of a mixture with the halide. Since lead metal is always produced from the halides, it is presumed that in a thin layer on a spark plug this can be oxidized to the necessary oxide-halide combination for further reaction with the phosphorus compound.

High-Temperature X-Ray Diffraction Studies — A study of the properties and formation of representative lead compounds was carried out by means of high-temperature X-ray diffraction. The high-temperature X-ray patterns taken of mixtures of pure lead oxides and pure lead halides revealed many chemical and physical changes at the temperatures characteristic of a firing spark plug. In some cases it was possible to estimate the temperatures of formation of new compounds such as oxyhalides. Many of these chemical reactions explain the occurence of certain compounds in spark-plug deposits. The results of this study have been reported in greater detail previously.[2]

The most important result from this work is the evidence for the irreversibility of the chemical reactions involved in the formation of spark-plug deposits. This study showed that the lead compounds from spark-plug deposits identified by X-ray diffraction analysis at room temperature are the same compounds present on the spark plugs during (high-temperature) engine operation. No chemical changes take place when the deposits are cooled to room temperature. Thus, the lead oxyhalides formed in the combustion chamber do not revert to lead oxide and lead halides at temperatures below the temperature of formation. Temperatures of formation of some commonly occurring compounds are shown in Fig. 3.

The frequent presence of PbClBr in deposits results from the low temperature of formation shown in Equation 1. From Equation 2, it may be inferred that the very similar $PbO \cdot PbClBr$ is found much less often than $2PbO \cdot PbClBr$ because the monobasic compound exists only in the approximate temperature range 550–900 F. Equation 4 accounts for the very common occurrence of $2PbO \cdot PbClBr$, which is stable from 670 to over 1300 F. This range of temperatures represents normal spark-plug operating temperatures.

Temperature, thus, is important in explaining differences in deposit composition.

Electrical Conductivity of Lead Compounds

Laboratory Test to Measure Conductivity of Lead Compounds — The electrical conductivity of lead compounds found in spark-plug deposits has been measured in a laboratory test. A diagram of the test apparatus is shown in Fig. 4.

The test procedure consists of heating the pow-

[2] "High Temperature X-Ray Study of Lead Halides and Oxyhalides," by H. P. Julien and R. E. Ogilvie. Paper presented at meeting of Division of Physical Chemistry, American Chemical Society, Boston, April, 1959.

Table 4 — Conduction Temperatures of Lead Compounds

	Conduction Temperature, F	Melting Point, F
Lead Oxybromides		
$PbBr_2$	530	700
$PbO \cdot PbBr_2$	650	925
$2PbO \cdot PbBr_2$	980	1310
Lead Oxyphosphates		
$Pb_3(PO_4)_2$	1260	1860
$PbO \cdot Pb_3(PO_4)_2$	1240	1800
$2PbO \cdot Pb_3(PO_4)_2$	1180	—
$5PbO \cdot Pb_3(PO_4)_2$	1140	1580
Lead Oxysulfates		
$PbO \cdot PbSO_4$	1400	1790
$2PbO \cdot PbSO_4$	1340	1740
$4PbO \cdot PbSO_4$	1290	1640
$PbSO_4$	1250	1590[a]

[a] Transition temperature.

dered sample — pure lead compound, complex lead salt, or mixture of lead compounds — until it becomes so conductive that the spark plug no longer fires. At room temperature the sample has high electrical resistance; hence, the plug fires normally. As the temperature rises, the resistance of the lead compound decreases and it begins to steal current from the spark plug. Finally, the plug is completely shorted out. Then the burner is removed. As the sample cools, its resistance increases. At some temperature the first spark reappears at the plug. This temperature is defined as the conduction temperature. Repeated heating and cooling establish this temperature within limits of ± 30 deg.

Role of Gaseous Ions in Misfiring — In one experiment, small piles of a spark-plug deposit were placed at intervals along the bottom of the ceramic sample boat. Upon heating, the sample became conductive, gaseous ions were apparently formed in the surrounding air, and the resulting continuous electrical path shorted out the spark plug. This behavior corresponds to that of the discontinuous deposits often observed on spark plugs.[3]

Correlation of Conduction Temperature with Melting Point — The conduction temperatures of many lead compounds correlate extremely well with their melting points. This is illustrated in Table 4 for the PbO–$PbBr_2$ system. The higher oxybromides have higher conduction temperatures and hence higher electrical resistance.

On the other hand, increasing the lead oxide content of the high-melting lead oxysulfates and oxyphosphates decreases the conduction temperatures of the complexes, since the electrical resistance of lead oxide is lower than that of lead sulfate or lead phosphate.

An interesting result is obtained with the lead oxysulfates. The conduction temperature of the simple lead sulfate does not fall in line with its melting point (1990 F). Rather, the conduction temperature is related to the transition temperature between the two crystalline modifications of $PbSO_4$.

Thus, melting point alone is a satisfactory indicator of conduction temperature only in the absence of physical and chemical changes, such as the phase transition mentioned above.

[3] "Spark Plug Fouling Studies," by R. J. Greenshields. Paper presented at SAE National Passenger-Car, Body, and Materials Meeting, Detroit, March, 1952.

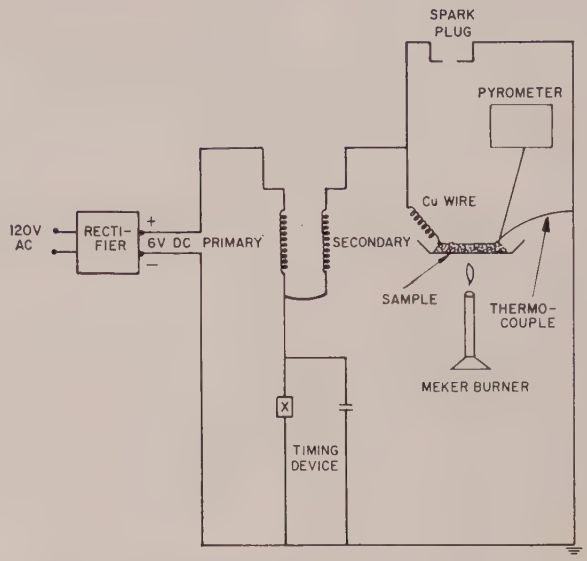


Fig. 4 — Laboratory apparatus for comparing resistance of lead compounds

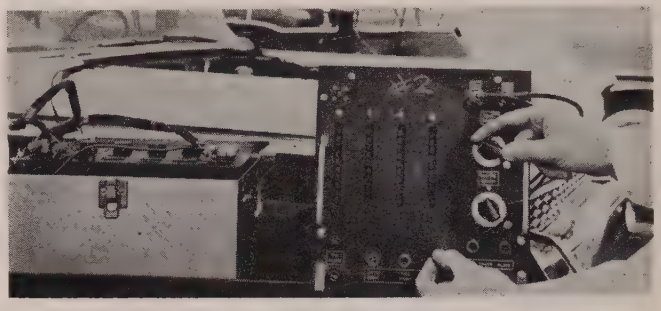


Fig. 5 — Spark-plug misfire counter

Engine Studies

The dependence of spark-plug deposit composition on temperature demonstrated in laboratory studies has been further developed in engine studies which related changes in misfiring severity to changes in spark-plug temperature and deposit composition. Unless otherwise indicated, these engine studies were performed in a laboratory V-8 engine of 10/1 compression ratio. Misfiring was measured at 80 mph, wide-open throttle. Throughout these engine studies misfiring has been measured quantitatively with a new electronic misfire counter.

Spark-Plug Misfire Counter

This instrument has been developed to measure quantitatively the number of misfires occurring in any or all cylinders of an engine under actual operating conditions. It can be applied to any spark-fired engine. Because it is designed to measure misfires due only to spark-plug malfunction, it has been of great value in investigations of fouling.

Misfire counters for both laboratory and mobile operation on the road have been developed and applied. The unit for mobile operation (Fig. 5) contains all components including the power sup-

plies, count register, and circuitry in a portable case. Connecting the unit to the battery and ignition system and making initial adjustments can be accomplished in a few minutes. No special or permanent attachments to the car are required. Cars equipped with either 6- or 12-volt batteries can be monitored by utilizing the appropriate misfire unit power supply.

For laboratory applications, the units are of the same basic design, but are equipped with a power supply to operate from the 110-volt ac line.

Further details of the instrument are given in the Appendix.

Correlation with Driver Reaction — As might be expected, the misfire counter can detect misfiring more easily than can the drivers of the vehicles. Nevertheless, the correlation between driver observations and misfiring severity as detected by the instrument is quite good. As is shown in Fig. 6, trained drivers do not begin to detect misfire until it reaches about 10–12%.

Spark-Plug Temperature during Deposit Buildup

Spark-Plug Deposit Formation at Different Temperatures — Two series of engine tests were carried out in the laboratory to evaluate the effect of spark-plug temperature during deposit buildup on deposit composition and misfiring. Deposits were accumulated with a commercial-type super-premium fuel containing no deposit modifier and a commercial 10W-30 oil. In the first series, deposits were formed at constant spark-plug tip temperatures of 600, 1000, and 1300 F. These temperatures were achieved by road-load operation at 30, 60, and 80 mph, respectively. The effect of a phosphorus additive was investigated at 600 and 1000 F. Engine conditions to give the desired temperatures were selected on the basis of an extensive study of the effect of engine operating conditions on spark-plug temperature in this engine. Misfire ratings were made by counting the misfires occurring during a 30-sec period at 80 mph, wide-open throttle.

At 600 F only PbClBr was formed with the base fuel. Misfiring increased rapidly to 50% at 28 hr. The misfire pattern was nearly the same on the 1000 F deposit of $2PbO \cdot PbClBr$. Obviously, the shunt resistances of these two compounds fell below the minimum value required to prevent misfiring at the temperature (1500 F) of the rating. However, the $PbO \cdot PbSO_4$ formed at 1300 F gave only negligible misfiring, as shown in Fig. 7.

Throughout these tests, which are summarized in Table 5, the differences in deposit composition were greater on the spark-plug tip than on its base, reflecting the greater change in temperature at the point farthest from the coolant.

When 0.2 theory of a phosphorus compound was added to the base fuel, misfiring was less at both 600 F and 1000 F than it was on the base fuel. (See Table 6.) Nevertheless, misfiring in the presence of deposits accumulated on phosphorus showed the same sensitivity to spark-plug temperature that the base fuel did. There was more misfiring on the 600 F deposits than on the 1000 F deposits. More misfiring on the low-temperature deposits was due to less complete conversion of the lead halides to the innocuous phosphorus complex at the lower temperature.

The results from the tests on the phosphorus additive are compared with those on the base fuel in Fig. 8.

The importance of the maximum spark-plug temperature attained during the formation of deposits was indicated by a series of tests run at 1000 F average spark-plug tip temperature. The 1000 F average temperature was achieved in four different ways: (1) road load operation at 60 mph, (2) 40 mph at three times road load, (3) 40 mph, road load, with the spark advanced 16 deg, and (4) 30 mph at four times road load.

Table 5 — Spark-Plug Temperature Affects Deposit Composition

Spark-Plug Tip Temperature during Deposit Accumulation, F	600	1000	1300
Equivalent Speed, mph	30	60	80
Spark-Plug Area	**Deposit Composition**		
Tip	PbClBr	2PbO · PbClBr ($PbO \cdot PbSO_4$)[a]	$PbO \cdot PbSO_4$ 2PbO · PbClBr
Base	PbClBr	PbClBr (2PbO · PbClBr)	PbClBr (2PbO · PbClBr)

[a] Parentheses denote minor constituent.

Table 6 — Low-Temperature Deposits Misfire More than High-Temperature Deposits

Spark Plug Tip Temperature, F	600	1000
Spark-Plug Area	**Spark-Plug Deposit Composition**	
Tip	$3Pb_3(PO_4)_2 \cdot PbClBr$ PbClBr	$3Pb_3(PO_4)_2 \cdot PbClBr$ PbClBr
Middle	$3Pb_3(PO_4)_2 \cdot PbClBr$ PbClBr	$3Pb_3(PO_4)_2 \cdot PbClBr$ PbO · PbClBr
Base	PbClBr $3Pb_3(PO_4)_2 \cdot PbClBr$	PbClBr $3Pb_3(PO_4)_2 \cdot PbClBr$
% Misfiring After		
20 hr	2	1
50	27	2
80	32	7
130	—	35

Table 7 — 1000 F Deposits Show Differences in Composition

Conditions		
Equivalent Speed, mph	Load	Spark Timing
60	Road	Standard
40	3 times road	Standard
40	Road	16 deg Advanced
30	4 times road	Standard
Deposits		
Tip	Middle	Base
2PbO · PbClBr		PbClBr
$PbO \cdot PbSO_4$[a]		2PbO . PbClBra
β-PbO	2PbO · PbClBr	PbClBr
$PbO \cdot PbSO_4$[a]		
$PbO \cdot PbSO_4$		PbClBr
2PbO · PbClBr		
NiO	2PbO · PbClBr	PbClBr
$PbO \cdot PbSO_4$	PbClBr	
2PbO · PbClBr		

[a] A minor constituent.

Significant differences in both deposit composition and misfiring were observed in this series of tests. Deposits formed at 60 mph were chiefly $2PbO \cdot PbClBr$ and some $PbO \cdot PbSO_4$. Deposits built up at 40 mph and advanced spark contained more $PbO \cdot PbSO_4$. This was due to the excessive exposure of the deposits to very high temperatures at the advanced spark setting. Although the indicated tip temperature was 1000 F, substantially higher temperatures apparently occurred during a part of each cycle.

The tip deposits from the run at 40 mph and heavy load contained lead oxide and various lead oxysulfates. The appearance of the PbO and the higher lead oxysulfate suggests the existence of higher temperatures. Furthermore, the indicated horsepower delivered by the engine was the same in this run as in the 60-mph, road load run. Since more power per stroke was being delivered at the lower speed, it is probable that higher maximum temperatures were attained.

Run 4 confirms the above conclusion. Here the speed was reduced to 30 mph and the load increased to maintain the 1000 F average tip temperature. X-ray analysis revealed NiO in the tip deposit. NiO originates from the center electrode at very high temperatures and is evidence for maximum temperatures even higher than at 40 mph. The composition of the deposits formed in these tests is summarized in Table 7.

The amount of misfiring observed in these tests was consistent with the deposit composition. Runs of 60 mph at road load and 40 mph at three times road load gave approximately the same misfiring since the predominant compounds in the deposits were those of relatively low resistance, $2PbO \cdot PbClBr$ and PbO. In the runs at advanced spark and four times road load, compounds of higher resistance predominated and misfiring was less (Fig. 9).

From the results of the tests at different average temperatures and those at a constant average temperature obtained in different ways, it can be concluded that spark-plug temperature is important in determining deposit composition but that the average temperature measured under a set of engine operating conditions is not sufficiently definitive. Because of the irreversibility of the chemical reactions in which the deposits are formed, the peak temperatures encountered are of greater consequence.

Effect of Plug Temperature on Shunt Resistance and Misfiring

In addition to determining the composition of deposits, spark-plug temperature also determines the shunt resistance of the spark plug. This effect is particularly important during accelerations, since the high temperatures which increase deposit electrical conductivity are attained quickly while deposit composition changes more slowly.

When knock occurs, spark-plug temperature is increased.[4] Under this condition more misfiring would be expected and has been found. A study (Fig. 10) in a 1958 model car showed that misfiring is more severe when knock occurs and that it in-

[4] "Spark Plug Fouling." E. I. du Pont de Nemours & Co.

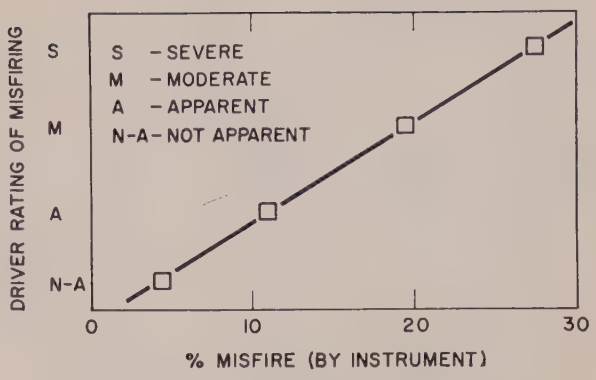


Fig. 6 — At about 10-12% misfire becomes apparent to trained observer

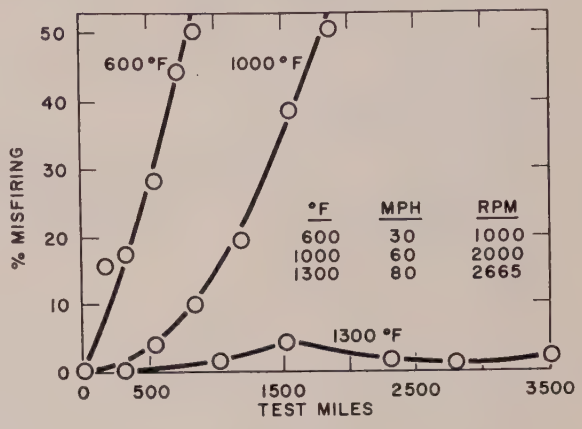


Fig. 7 — Misfiring decreases with increasing temperature of deposit formation

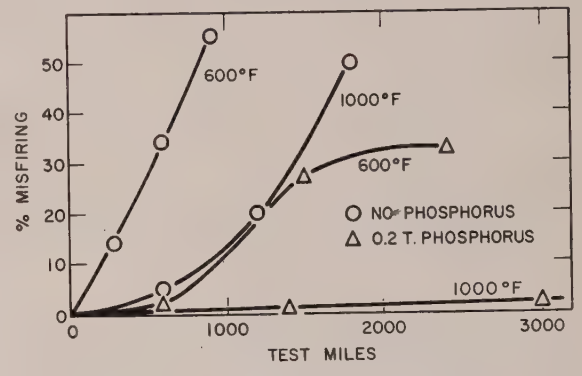


Fig. 8 — Temperature of deposit accumulation affects misfiring on phosphorus and nonphosphorus fuels

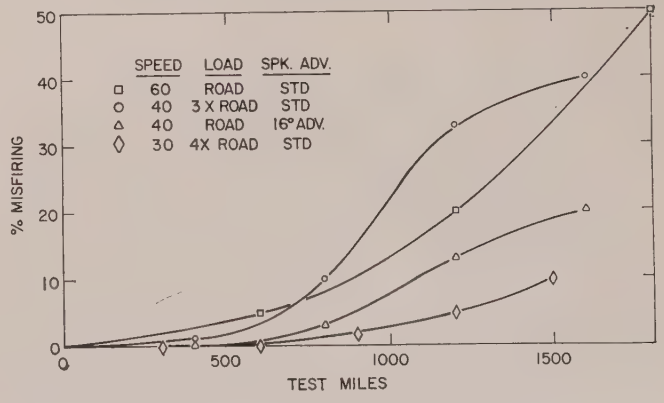


Fig. 9 — 1000 F deposits show differences in misfiring

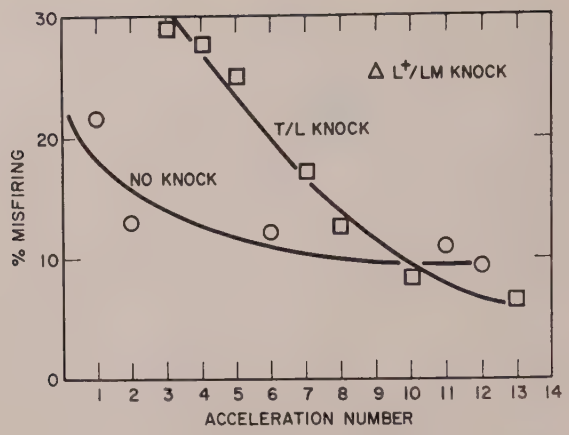


Fig. 10 — Misfiring is increased by knock

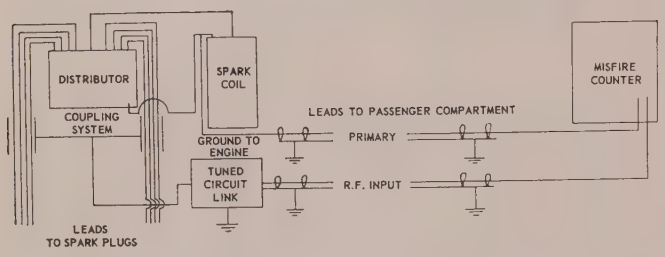


Fig. 11 — Method for coupling signal from ignition system to spark-plug misfire counter, block diagram

creases in severity when knock intensity increases.

This effect is of considerable importance in evaluating cars for misfiring. Care should be taken to ensure that the fuel quality is sufficient to prevent knock.

One other observation in a passenger car on a test strip showed that the effect of temperature on shunt resistance and misfiring is exercised very rapidly. Temperatures observed at the end of accelerations made from 28–80 mph and 50–80 mph were approximately the same, the latter being about 50 deg higher. Thus, the amount of misfiring experienced in the two accelerations would be expected to be, and was found to be, the same.

These data demonstrated that at the lower speeds spark-plug temperature was too low to cause an appreciable amount of misfiring, and that above 50 mph approximately equal amounts of misfiring were observed since the same temperatures were attained in both accelerations (Table 8).

Conclusions

Those factors which change deposit composition during deposit accumulation or spark-plug temperature during misfire ratings will exert a significant effect on the severity of misfiring.

Severity of misfiring is increased by higher spark-plug temperatures during misfire evaluations. This increase is effected by running at higher speed, load, spark advance, or any other engine operating condition which might be expected to give higher combustion chamber temperatures. For example, when knock occurs, temperatures are higher and misfiring is more severe than in the absence of knock. The effect of the higher temperatures is to decrease the resistance of the compounds present on the spark-plug ceramic.

Misfiring severity under a given set of rating conditions is also increased by forming deposits at lower spark-plug temperatures. Compounds formed at lower temperatures have relatively low melting points and low electrical resistance; consequently, they lead to more misfiring. The melting point of a compound, provided there are no changes in its physical structure on heating, may be used as an indicator of its electrical conductivity, and thus of its misfiring tendency.

Conversion of spark-plug deposits to compounds having higher melting points and higher electrical resistance reduces misfiring. This change may be accomplished by the use of phosphorus additives which form deposits of high resistance even under conditions of low-temperature operation. The presence of lead oxide seems essential for this conversion. Deposits of high electrical resistance are also formed under conditions of high-temperature operation. However, prolonged operation at high temperatures is not essential. Because the reactions leading to the conversion of low-temperature deposits to high-temperature deposits are irreversible, relatively brief exposures to high temperature are sufficient to effect at least a partial conversion.

Acknowledgment

The authors wish to acknowledge the participation of their many coworkers in the research leading to this report. In particular, we wish to mention J. J. Heigl and J. A. Wilson for the development of the spark-plug misfire counter. The contributions of R. H. Clinard, F. C. Gunderloy, Jr., L. B. Shore, F. E. Steidler, and the late J. E. Neudeck are also gratefully acknowledged.

APPENDIX

Spark-Plug Misfire Counter

Operating Principle — One major component of the radio frequency signal emanating from the ignition system is a function of the firing characteristics of the spark-plug. There is a minor component related to the distributor rotor arc, but this can be minimized so that the detected signal is a primary function of the spark-plug arcing action. Two types of condition at the plug reduce the radio frequency signal: (1) absence of arc, (2) presence of an arc in a rich mixture as on deceleration. A lean mixture appears to produce an erratic signal, which varies in intensity from cycle to cycle. On normal firing, as during constant speed running or during an acceleration, the radio frequency signal is constant.

In the misfire counter, the radio frequency signal along with an on-off voltage produced by the breaker-point action is utilized. When the spark fires normally, a satisfactory radio frequency signal coincides with the breaker-point signal and the electronic circuit does not allow the register to operate. However, when the spark-plug misfires, no radio frequency signal is produced at the plug and the breaker-point signal produces a count on the register. Each misfire registers as an additional count.

A vital part of the circuit is the time delay se-

quence. The radio frequency signal prevents the counter from registering for 3 millisec after the spark. During this interval, the signal from the primary of the spark coil (activated by the breaker point) is attempting to operate the counter. The primary coil signal attempts to operate the counter 1.5 millisec after the distributor breaker points open.

Through a switching arrangement, the radio frequency signal is made inoperative when total count, equivalent to the total number of firing opportunities, is to be measured. This is done by counting the total number of times the primary coil is activated by the breaker points.

Description of the Instrument — The radio frequency and the primary voltage signals are coupled, as shown in Fig. 11, from the high-tension leads to the tuned input circuit and from the primary of the spark coil, respectively. The selected frequency to which the circuit is tuned is 1.5 megacycles. The unit designated as the "tuned circuit link" is contained in a small, shielded box which is coupled with short (approximately 6-in.) leads to the high-tension leads and to a ground (engine frame). This unit is connected through a lead of any convenient length to the tuned circuit, which is located on the chassis of the instrument. A block diagram shows the functions of the electronic elements (Fig. 12).

Application — The misfire counter can be applied to any spark-fired engine. The results obtained may be expressed as the percentage misfire, average number of misfiring plugs, or total number of misfire events. In order to measure percentage misfire, both the number of firing opportunities and the number of misfire events are counted. The counts are displayed on a register to four figures. The unit is also equipped with a switching circuit, so that both firing opportunities and misfire count can be obtained on consecutive runs. It is possible during a normal vehicle acceleration to record the indicated values for each 10-mph speed increment.

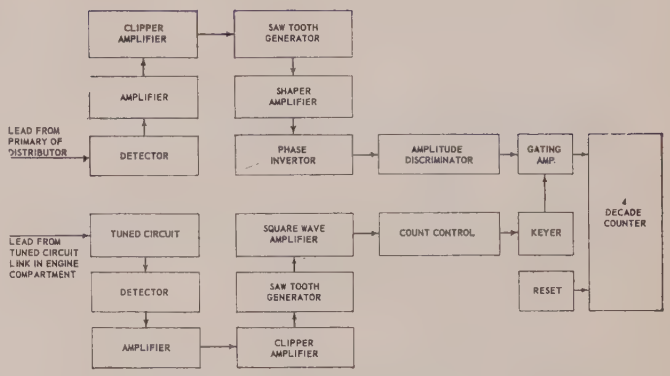


Fig. 12 — Spark-plug misfire counter, block diagram showing functions of electronic elements

Table 8 — Misfiring at Various Accelerations

	Cumulative Misfire Counts					
Mph	30	40	50	60	70	80
20–80 mph Wide-Open Throttle Acceleration	0	0	45	340	1120	2910
50–80 mph Wide-Open Throttle Acceleration	—	—	—	205	910	2940

Additional References

1. "Spark Plug Problems in Automotive Service," by R. C. Beaubier, H. J. Chalk, and M. M. Roensch. Paper presented at SAE National Passenger-Car, Body, and Materials Meeting, Detroit, March, 1952.

2. "Factors Affecting Functioning of Spark Plugs," by W. A. Bychinsky. SAE *Quarterly* Transactions, Vol. 2, April, 1948, pp. 255–267.

3. "Studies in Lead Chloride-Lead Bromide System," by G. Calingaert, F. W. Lamb, and F. Meyer. *Journal* of American Chemical Society, Vol. 71, 1949, p. 3709.

4. "Engine Laboratory Study of Motor Oil Characteristics as One of Factors Influencing Combustion Chamber Deposits," by D. E. Carr, L. A. McReynolds, S. C. Britton, and R. E. Linnard. Paper presented at SAE Annual Meeting, Detroit, January, 1952.

5. "Some Relationships of Fouling to Spark Plug Design and Operation," by R. K. Christie. Paper presented at CRC meeting, Dearborn, Mich., September, 1952.

6. "Control of Combustion Chamber Deposits," by C. L. Fleming, N. V. Hakala, L. E. Moody, and C. O. Tongberg. Paper presented at National Petroleum Association meeting, Cleveland, April, 1954.

7. "Spark Plug Fouling — Survey — Test Procedures — Fuel Factors," by C. A. Hall, R. C. Beaubier, E. C. Marckwardt, and R. L. Courtney. Paper presented at SAE Annual Meeting, Detroit, January, 1957.

8. "Improved Gasoline Performance with Phosphorus Additives," by R. E. Jeffrey, L. W. Griffith, and E. Dunning. Paper presented at National Petroleum Association meeting, Cleveland, April, 1954.

9. "Phase Study of PbO-$PbBr_2$ System by X-Ray Diffraction," by F. W. Lamb and L. M. Niebylski. *Journal* of American Chemical Society, Vol. 75, 1953, p. 511.

10. "Formation of Engine Deposit Compounds by Solid-State Reactions (X-Ray Diffraction Study)," by F. W. Lamb and L. M. Niebylski. Paper presented at meeting of API Division of Refining, Tulsa, April, 1951.

11. "Fundamental Studies Related to Spark Plug Deposits," by J. L. Lauer, P. J. Fried, and R. D. Miller. Paper presented at SAE National Fuels and Lubricants Meeting, Tulsa, November, 1958.

12. "Spark Plug Fouling from Car Manufacturer's Viewpoint," by H. N. Metzel. Paper presented at SAE National Passenger-Car, Body, and Materials Meeting, Detroit, March, 1952.

13. "Fuel Deposits and Spark Plug Fouling," by T. B. Rendel. Paper presented at NPA meeting, Cleveland, April, 1952.

14. "Gasoline Additives," by C. O. Tongberg, N. V. Hakala, L. E. Moody, J. B. Patberg. Paper presented at SAE Annual Meeting, Detroit, January, 1954.

15. "Operation of Spark Plugs in Present Day Engines," by H. H. Vogel. Paper presented at SAE National Passenger-Car, Body, and Materials, Meeting, Detroit, March, 1952.

16. "Aviation Spark Plug Fouling — Its Cause and Control," by V. E. Yust and E. A. Droegemuller. Paper presented at SAE Annual Meeting, Detroit, January, 1952.

Brake investment

CONSIDERATIONS in brake and brake drum rating, design, and testing are:

1. An engineering term to define brake, brake lining, and brake drum heat dissipating ability is necessary. This paper describes one possible term — continuous horsepower.

2. Brake lining material is a current weak link. Unless a major breakthrough in this area is forthcoming, no major strides in reducing both the weight and cost of brake and drum assemblies appears possible.

3. Under the limitations imposed by current materials, the major element not yet sufficiently appreciated by the industry is drum positioning and exposure.*

IN THIS PAPER we will cover in detail only those elements of brake design involved in drum design and testing. However, this coverage must, of necessity, include elements not normally considered in drum design alone, since the drum is only part of a braking system and cannot be divorced therefrom.

Rating of Brakes and Brake Drums

Since a vehicle brake is a mechanical device for converting the kinetic energy of the moving vehicle into heat energy and subsequently disposing of this heat, each element of the brake system must be designed for a particular capacity. This capacity must be based on the weight and expected speed of the vehicle as a whole, as well as the grades that it will negotiate. Unfortunately, up to the present time, there has been no accepted term available with which one could describe this capacity. As a result of this inability to express braking capacity, most manufacturers of vehicles have developed test procedures which will qualitatively prove the acceptability of their brakes on a particular vehicle. However, a minor change in, or relocation of, any of the component parts may have a major effect on the ability of the brake system to convert the kinetic energy into heat energy or subsequently to dispose of the heat.

The accepted terms used by the industry to describe brake effectiveness or performance are appropriate for all facets of the problem, except that of fade resistance. Brake fade or loss of effectiveness occurs when heat is generated at a faster rate than it can be dissipated. The problem of properly describing fade resistance, therefore, becomes one of semantics.

About 10 years ago our company decided to describe fade resistance in terms of continuous or steady state horsepower. Our decision was based on the obvious fact that all automotive engineers are familiar with the term "horsepower" as used to describe engine power and that the brake capacity required must be in a certain ratio with the power available in the vehicle engine. The brakes should not be expected to carry the vehicle down a grade or to stop it from a speed which is beyond the capabilities of the vehicle engine. The words "steady state" or "continuous" separate the problem of heat storage from that of heat dissipation. Under steady state conditions, heat is dissipated at the same rate that it is generated; and a brake, operating at its peak capacity under steady state conditions, will theoretically operate indefinitely — the only type of failure possible being due to accelerated brake lining wear.

With the adoption of our added descriptive term for brake fade resistance, we had a means of comparing brakes without reference to size or type. We could pinpoint advantages of different designs of drums and of different drum materials. Evaluation of brake linings became more simple. However, it must be borne in mind that the new method of test developed was used in addition to all other required tests.

The need for the new term becomes more obvious when we find that the Federal Government, in submitting Invitations to Bid on new military vehicles, gives braking requirements in such terms as "Brakes must be capable of stopping the unit on a 30% grade from 40 mph *without fade,*" and "Vehicle must be

* Paper presented at SAE National Transportation Meeting, Chicago, Oct. 27, 1959.

Drums—

for safety and profit

Paul G. Hykes and Clarence A. Herman

The Budd Co.

able to stop in 128 ft from 40 mph *consistently*." The italics are ours to emphasize that, in our opinion, the wording is ambiguous and could be most misleading.

The present Interstate Commerce Commission regulation on the performance requirements for brakes states:

> "193.52 *Brake Performance.* Every motor vehicle or combination of motor vehicles shall, upon application of the service brakes, be capable, at all times and under all conditions of loading, of being brought to a stop within a braking distance of 30 ft from a speed of 20 mph when tested on a dry, smooth, level road free from loose material. The braking distance shall be measured by means of an instrument or machine of the decelerometer type capable of being read in feet to stop from 20 mph."

You will note that a strict interpretation of this rule would cover conditions of brake fade, since a truck or combination vehicle could be required to pass this test immediately after descending a long mountain grade. Perhaps, for this reason, no other requirements are listed on this subject.

The Uniform Vehicle Code (UVC) has similar requirements. However, since measured stopping distance is required rather than decelerometer readings, an allowance for time delay in actuating the brakes is permitted. Here too, the brakes can be called upon to perform as required at any time.

In 1956 the Commonwealth of Pennsylvania, noting that certain accidents were being caused by vehicles which were either underbraked or overloaded, passed an ordinance that required all vehicles to have at least 1 sq in. of brake lining for each 55 lb of gross vehicle weight. The Automobile Manufacturers Association and the Truck-Trailer Manufacturers Association viewed this activity with alarm, since it was a step toward design by legislation which in the long run would stifle engineering advances. Accordingly, the problem faced by Pennsylvania, and obviously other States, in promoting highway safety through legislation was referred to the Joint AMA-TTMA Committee on Vehicle Brakes. The Committee was asked to develop a method whereby the brakes on commercial vehicles could be realistically rated and certified by the manufacturer as to capacity.

The AMA-TTMA Brake Committee is considering the Budd test method and rating system. With minor revisions this method is now undergoing tests and results to date look promising. Meanwhile, the SAE Brake Committee is reviewing the overall problem, and it appears reasonable to assume that in the near future a new term will be added to our brake engineering vocabulary.

Thermal and Power Considerations

Before going into the problems of designing brake drums, let us first study some of the conditions un-

Table 1 — Parasitic versus Brake Horsepower Required for Various Vehicles on Grades

Conditions: 3500 rpm engine speed
Frontal area—75.00 ft² for 20,000–50,000 gvw
50.83 ft² for 10,000 gvw
39.00 ft² for 5000 gvw

Gvw, lb	Vehicle Velocity, mph	Parasitic Horsepower Due to				Total Bhp	Grade Required to Maintain Speed, %	Grade Permitted with Brakes, %
		Rolling Resistance	Air Resistance	Chassis Friction	Total Parasitic			
5000	20	2.51	1.78	4.50	8.79	6.5	3.30	5.73
	30	4.12	6.07	4.50	14.69	6.5	3.67	5.30
	40	5.98	14.28	4.50	24.76	6.5	4.64	5.86
10,000	20	5.01	2.33	8.00	15.34	13.0	2.88	3.44
	30	8.24	7.92	8.00	24.16	13.0	3.02	4.65
	40	11.95	18.62	8.00	38.57	13.0	3.62	4.83
20,000	20	10.02	3.59	15.05	28.66	26.0	2.69	5.12
	30	16.48	12.20	15.05	43.73	26.0	2.73	4.36
	40	23.90	28.71	15.05	67.66	26.0	3.17	4.39
30,000	20	15.03	3.59	22.00	40.62	39.0	2.54	4.98
	30	24.72	12.20	22.00	58.92	39.0	2.46	4.08
	40	35.85	28.71	22.00	86.56	39.0	2.71	3.92
40,000	20	20.04	3.59	29.05	52.68	52.0	2.47	4.91
	30	32.96	12.20	29.05	74.21	52.0	2.32	3.94
	40	47.80	28.71	29.05	105.56	52.0	2.47	3.69
50,000	20	25.05	3.59	36.00	64.64	65.0	2.42	4.86
	30	41.20	12.20	36.00	89.40	65.0	2.24	3.86
	40	59.75	28.71	36.00	124.46	65.0	2.33	3.55

NOTE: Computations based on SAE truck ability prediction procedure TR-82 for parasitic hp.

der which the brakes and drums convert energy into heat. Since most fleet operators are familiar with the drum for the popular 16½ in. × 7 in. over-the-highway truck and trailer brake, it might be best to concentrate on that size. The drum is used on vehicles having a legal weight limit of 18,000 lb per axle and many of the vehicles are capable of speeds of 70 mph.

In stopping such a vehicle from 70 mph, one brake will absorb the energy of 9000 lb. This energy is simple to calculate:

$$KE = \tfrac{1}{2} MV^2 = \frac{WV^2}{2g}$$

$$= \frac{9000 \times 102.67^2}{2 \times 32.2} = \frac{9000 \times 10{,}500}{64.4} \qquad (1)$$

$$= 1{,}470{,}000 \text{ ft-lb}$$

where:

KE = Kinetic energy
M = Mass
V = Velocity, fps
W = Weight, lb
g = Acceleration due to gravity = 32.2 fps²

Since one Btu is equivalent to 778 ft-lb, the total heat generated in one stop is:

$$Q_T = \frac{1{,}470{,}000 \text{ ft-lb}}{778 \text{ ft-lb/Btu}} \qquad (2)$$

$$Q_T = 1{,}895 \text{ Btu}$$

If the stop were made according to the UVC requirements of braking force equal to 43.5% of gross weight, the stopping rate is 0.435×32.2 fps² = 14 fps² and the actual stopping time would be $\frac{102.67 \text{ fps}}{14.0 \text{ fps}^2}$ or 7.33 sec.

Thus, the 1895 Btu total heat generated in the stop would be absorbed in 7.33 sec or $\frac{1{,}895 \text{ Btu}}{7.33 \text{ sec}}$ = 258 Btu/sec. Since 1 hp = 0.707 Btu/sec, the average horsepower in the stop is $\frac{258 \text{ Btu/sec}}{0.707 \text{ Btu/sec hp}}$ = 365 hp.

This is the average horsepower developed during the complete stop. At the beginning of the stop, the peak horsepower is twice the average power, since the instantaneous power varies directly with the speed, and thus drops off to zero as the speed reaches zero. Another way of figuring the peak power (the instantaneous power at the beginning of the stop) is from the braking force and speed. In our stop the peak power is:

$$\text{Peak Power} = \frac{F\,V}{550} = \frac{9000 \text{ lb} \times 0.435 \times 102.67 \text{ fps}}{550 \text{ ft-lb/sec/hp}}$$

$$= 730 \text{ hp} \qquad (3)$$

This is twice the *average* stopping power.

Perhaps a better formula for the automotive engineer's use in calculating instantaneous power from braking force and speed is:

$$\text{Power} = \frac{F\,V}{375} \qquad (4)$$

where:

F = Force, lb
V = Velocity, mph

In our case:

$$\text{Power} = \frac{9000 \text{ lb} \times 0.435 \times 70 \text{ mph}}{375 \text{ mile-lb/hr/hp}}$$

$$= 730 \text{ hp}$$

Fortunately, a truck or trailer combination is never called upon to develop braking power continuously at these highpower rates because:

1. On level roads, more than 730 hp per brake

would be required to keep the speed at 70 mph.

2. In descending grades, the brakes are usually snubbed, and the average power is the same as if the brakes had been applied continuously. To develop this continuous peak at 70 mph with dragging brakes would require a grade angle the sine of which is 0.435. This is an angle of 25¾ deg or a grade of 48.3%. Such a grade is out of the question on public highways and even if there were such a grade, it would probably be too short to have the power at the peak level very long and most likely it would be extremely unsafe to descend at 70 mph.

3. The power as calculated and shown is the total power exerted in stopping the vehicle; however, the brake does not convert all this power into heat. It is helped by wind resistance, tire rolling resistance, and friction in the rotating parts of the wheels and drive system components, all of which are otherwise known as parasitic power losses when referred to in their effect on reducing engine driving power. Engine braking also helps retard the vehicle.

Having considered peak braking power and average braking power, we will proceed to continuous power by again studying the vehicle making repeated stops on level roads from 70 mph.

Such a vehicle should be capable of making more than one stop per mile from 70 mph, but let's assume that one stop per mile is all that the engine will permit. In making the stop, the brake drums will absorb the heat and prior to the next stop, this heat must be dissipated. At an average speed of 35 mph there would be 103 sec for the heat to be dissipated. Therefore, the 1895 Btu should be dissipated at the

$$\text{rate of } \frac{1895 \text{ Btu}}{103 \text{ sec}} = 18.5 \text{ Btu/sec.}$$

We now have a power requirement which we can convert into chp (continuous horsepower), thus

$$\frac{18.5 \text{ Btu/sec}}{0.707 \text{ Btu/sec/hp}} = 26 \text{ chp.}$$

So far we have given no consideration to the parasitic losses such as rolling resistance, engine drag, or wind resistance; and when the 26 chp is compared to the actual capacity of brakes, as tested on a dynamometer, it is found to be at least 30% high.

We have found by extensive testing that the steady state capacity of a brake and drum assembly should be in the range of 1.2–1.3 chp per 1000-lb brake lead for relatively safe highway operation.

Fig. 1 shows graphically the instantaneous power during the stop, the peak power at the beginning of the stop, the average power during the stop, and the continous power in a series of stops. The difference between peakpower and continuous power of brake drums can be compared quite simply to the storage capacity of a funnel and its discharge rate. A funnel could have gallons of storage capacity and gallons per minute as a discharge rate. This would compare to Btu's storage capacity and Btu's per minute (or continuous horsepower) as the discharge or dissipation rate for the brake drum. As in the funnel, the two characteristics, storage capacity and discharge rate, are entirely divorced from each other.

Fig. 2 illustrates this point. The ship with a lot of power is apt to take on much water in a heavy sea. The designers have anticipated this and have provided a good bilge pump to eliminate the excess water. A certain amount of safety is assured. The

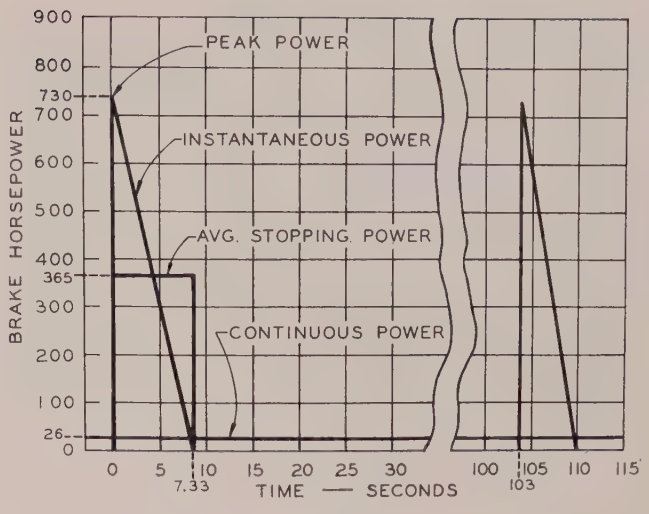


Fig. 1 — Braking power

Fig. 2 — Water storage capacity and water discharge rate

small boat that has lost its sail may be in a different predicament. Chances are that all the operator can do is to bail like a hero until it is safe to row to port. The analogy is relatively clear. Heat stored, like water taken above decks, must be got rid of or disastrous consequences are inevitable.

So far, we have analyzed fade resistance from the standpoint of high-speed stops only. This is only a portion of the problem, since a series of snubs from lower speeds can also build up destructive temperatures. Long mountain grades also must be considered. However, with the new term "continuous horsepower", we can predict the grade that a vehicle may safely negotiate. The formula is simple and easily understood:

$$\text{Grade} = \frac{\text{chp} \times 375}{\text{gvw} \times \text{speed}} \qquad (5)$$

where:

Grade = Per cent
chp = Continuous hp capacity of vehicle brakes
Speed = Mph

For actual application of this formula, the rolling resistance, wind resistance, chassis friction, and engine drag must be added to the chp of the brakes.

Table 1 shows data based on the SAE Truck-Ability Prediction SP–82. We are indebted to W. T. Burwell for all calculations, except those concerning

vehicle brakes.[1] The effect of brakes has been added on the basis of 1.3 hp/1000 lb gvw. This figure is, as pointed out earlier, one which appears possible. Note that these calculations cover grades of indefinite length; therefore, true steady state conditions must obtain.

We have shown that one of the requirements of a brake drum is that it be able to absorb heat at peak power to store it temporarily, and to dissipate it later at its continuous power. It is well-known that this heat can be stored in a heavy brake drum at a lower temperature than in a light drum. The problem then is a compromise between drum weight and temperature. The drum weight necessary to store the 1895 Btu developed in one stop can be determined from the formula:

$$W = \frac{Qs}{C_p(T_2 - T_1)} \qquad (6)$$

where:

W = Weight of drum, lb
Qs = Quantity of heat stored, Btu
C_p = Specific heat of drum material, Btu/lb/F (0.12 for cast iron)
T_1 and T_2 = Initial and final temperature of drum, F

As an example, if 200 F is a safe average drum operating temperature,

$$W = \frac{1895 \text{ Btu}}{0.12 \text{ Btu/lb/F } (200 - 70)}$$

$$W = 121 \text{ lb}$$

Similarly, if a safe average drum temperature is 300 F a drum weight of 69 lb will suffice.

However, if a vehicle is able to make stops from higher speeds than 70 mph, or if it is called upon to descend grades at higher speeds for short durations, the thermal storage capacity of the drums should be increased by additional weight to take care of the higher peak power demands.

Apparently, all persons concerned with heat problems agree that heat is dissipated by radiation, conduction, or convection. Applying these principles to brake assemblies, the heat can be:

1. Radiated to surrounding parts. This is undesirable. Tires build up temperatures on their own account.

2. Conducted to surrounding parts. This also is undesirable. The closest parts, such as the tire, the hub, and its bearings should not be overheated.

3. Convected into the passing air. Here is the place for the heat! While heat is dissipated by convection mostly from the drum ring, it also is convected from the surrounding parts which have been heated by radiation and conduction.

Since radiation might be a major element in brake heat dissipation, let's investigate a little, going back again to our standard 16½ in. × 7 in. drum braking a 9000-lb load to a stop from 70 mph. We previously calculated the heat from that stop to be 1895 Btu. We also calculated that a 69-lb drum would be heated to 300 F in absorbing 1895 Btu. Assume the drum ring has an exterior surface area of 3 sq ft, all at 300 F, and that the ambient temperature of the air and surrounding parts is 70 F. Then by an accepted formula (Stefan-Boltzmann Law) for radiation, we solve as follows:

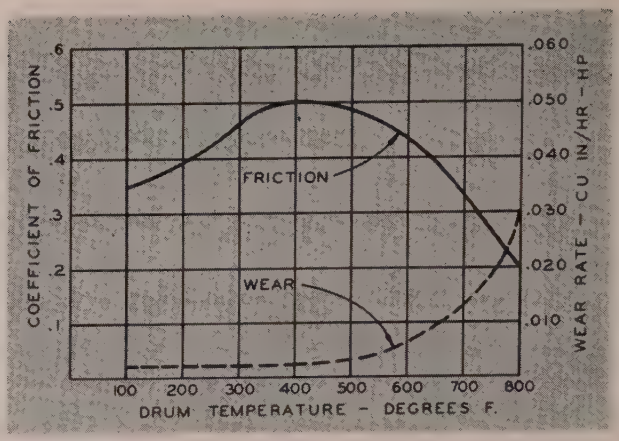


Fig. 3 — Friction versus temperature and wear versus temperature

$$Q_r = 1.73 \times 10^{-9} \times Z \times (T_D^4 - T_S^4) \qquad (7)$$

where:

Q_r = Rate of heat radiation, Btu/hp/ft²
1.73×10^{-9} = Stefan-Boltzmann constant in Btu/hr/ft²/R⁴
Z = Emissivity factor = 0.90 for cast iron
T_D = Drum temperature, R
T_S = Temperature of surroundings, R

$$Q_r = 1.73 \times 10^{-9} \times 0.90 \times (760^4 - 530^4)$$
$$= 400 \text{ Btu/hr/ft}^2$$

For our drum the total heat radiated is:

$$Q_t = 400 \text{ Btu/hr/ft}^2 \times 3 \text{ ft}^2$$
$$= 1200 \text{ Btu/hr}$$

It would, therefore, require more than 1½ hr to radiate the 1895 Btu if the drum surface temperature stayed at 300 F; but since its temperature drops as it loses heat, it would require more than 3 hr to radiate all of this heat. Obviously, radiation plays only a minor part in heat dissipation from brake drums.

Now to calculate the heat that might be conducted away from the drum. Brake lining in general, with the exception of the newer sintered metal lining, is a poor conductor so the heat carried into it and the brake shoes is negligible. The majority of the heat thus goes into the drum; and if it were dissipated entirely by conduction, it would be conducted from the drum ring through the drum head into the hub flange and then into the hub and wheel or wheels. This path of conduction is more or less devious and is interrupted by heat barriers at all contact points between the various parts.

However, for our purpose, we can assume the conduction through the contact points is the same as through cast iron. We can also assume the heat is conducted through the drum ring for an average distance of 6½ in., through its cross-sectional area of 25 sq in.

If we again assume the drum is heated to 300 F and the hub and wheels stay at 70 F, the total heat conducted from the drum into the wheels can be

[1] W. T. Burwell, Specialized Vehicle Engineering, Chevrolet Motor Division, General Motors Corp. Private communication.

calculated from the formula:

$$Q_c = K \times (T_1 - T_2) \times \frac{A}{L} \quad (8)$$

where:

Q_c = Quantity of heat conducted, Btu/sec
K = Thermal conductivity factor
= 0.00073 Btu/sec/in.^{2}/(F/in.) for cast
= iron
$T_1 - T_2$ = Temperature differential, F
A = Cross-sectional area, sq in.
L = Length of heat path, in.

Then:

$$Q_c = \frac{0.00073 \times (300 - 70) \times 25}{6.5}$$
$$= 0.65 \text{ Btu/sec}$$

Since it takes 1895 Btu to raise the temperature of the drum to 300 F it would take $\frac{1895 \text{ Btu}}{0.65 \text{ Btu/sec}}$ or 2920 sec or 49 min to conduct away the heat of one stop.

As with radiation, the actual time would be more than twice as long since the drum temperature would start at 300 F and immediately drop off while the wheel temperature would start at 70 F and increase, thereby decreasing the differential temperature.

It is also obvious that conduction, like radiation, is but a minor factor in helping to dissipate heat from the brake drum, so the only method left to dispose of the heat fast enough for the brake to perform satisfactorily is convection.

Brake Drum Design and Materials

The usual goals in designing current automotive products are maximum performance, long life, light weight, and low cost achieved through economy of manufacture and low material cost. At our company, the additional and more specific objective for brake drum design is to have it function at as low a temperature as possible. There are several reasons for this approach which are as follows:

1. The friction characteristics of contemporary organic type brake lining materials vary appreciably with temperature. Most are reasonably stable at temperatures approximating 500 F but lose much of their friction at 800 or 1000 F. This is the main cause of "fade."

2. The wear rate of organic lining increases rapidly as the operating temperature of the brake and drum increases. Fig. 3 shows how the friction and wear rate change with temperature.

3. Brake drums expand with increasing temperature, thus requiring more brake shoe travel and additional wheel cylinder piston displacement. If the air cylinder, wheel cylinder, or master cylinder bottoms, the deceleration falls off. In addition, the radius of curvature of the drum compared to that of the lining changes, producing decreased contact area and localized high temperatures. These are the secondary causes of "fade."

4. At elevated temperatures, the strength of any commercial brake drum material is drastically reduced. V. A. Crosby has shown tensile strengths lowered as much as 30% at 932 F from those at room temperature;[2] and, undoubtedly, brake drum irons are softer and have a greater tendency to score and wear.

5. At higher ambient temperatures in the drum and brake, the problems of vapor lock, corrosion, frozen bearings, softened springs, and even tire fires are more prevalent. It should be pointed out that any improvement in lining toward additional stability or lower wear might result in still better brakes. For example, a slight increase in the lining characteristics might permit designing a smaller brake which, because of better circulation of air around the drum, would offer greater benefits than can be attained through brake drum design only.

The factors of brake drum design and the physical and metallurgical characteristics of materials used in them contributing most toward the desired goals are listed with explanations of their individual purposes or functions:

1. Mass and Specific Heat — As covered in the section on power, the drum must have enough material in it to be able to absorb the heat generated in the peak power conditions of high-speed stops or snubs and store it at a low, safe operating temperature until it can be dissipated. Mass ties in directly with specific heat, since the product of mass and specific heat is thermal capacity.

2. Surface Area — The maximum surface area attainable is most important in lowering the brake drum temperature. Because most of the heat is dissipated from the brake drum by convection, the benefits gained from large surface area are quite obvious; and if the additional surface area can be arranged in ribs or fins to increase the flow of air over the drum, extra benefits will be derived. Finally, although the amount of heat dissipated by radiation to the surrounding parts is relatively small compared to the heat convected, additional surface area increases radiation.

3. Homogeneous Construction — This is desirable to avoid barriers to the flow of heat from the heat generating surface to the heat dissipating surface. This is particularly true of multipart drums having cast-iron braking surfaces mechanically attached to a carrier made of high conductivity metal such as aluminum. It is also true to a lesser extent of bimetallic drums having cast-iron liners supposedly bonded to high-conductivity metal where either the bond was defective when the drum was made or became so through use.

4. Exposure — Again, because the majority of brake heat is dissipated by convection, it is imperative that the drum, or as much of it as possible, be exposed to the air normally passing by the brake drum, wheel, and tire. Likewise, it helps if the extensions in the form of ribs, fins, or other shapes act as scoops or impellers to increase the circulation of air.

5. Shrouding — Shrouding is closely associated with exposure but is more directly connected with designing the drum to leave as much space for air to circulate between the drum and wheel rim as possible.

6. Strength — The strength designed into a brake drum is a combination of the physical dimensions and material characteristics. The need for strength in a drum quite obviously is to permit it to function properly for a long time under the severe mechani-

[2] "Metallurgical Developments in Brake Drums," by V. A. Crosby. SAE Transactions, Vol. 67, 1959, pp. 441–448.

cal and thermal stresses imposed. It must not only do this in its as-designed, original condition but in its worn or re-bored condition.

7. Rigidity — Rigidity, or its opposite, flexibility, must be designed into a brake drum to allow it to conform to brake-shoe deflection under load and still preserve, as nearly as possible, uniform pressure over the face of the lining.

The important remaining material characteristics controlling the performance and life of a brake drum are basic metallurgical characteristics or are dependent upon the metallurgical properties of the drum material. To cover the points we quote excerpts from a summary of the metallurgical requirements of cast iron for heavy-duty brake drums prepared by Earl Morrison.[3]

> "The following appears to be a fairly complete list of the basic requirements of brake drum iron to improve its service performance and minimize service failures.
>
> 8. High thermal conductivity, to minimize the attainment of damagingly high temperatures in a relatively thin layer of iron at the drum face in contact with the brake shoe lining. Excessive temperatures in this area can cause heat checking, high stresses due to severely localized thermal expansion, phase changes in the matrix and even melting in extreme cases. High thermal conductivity promotes higher temperatures on the outside surfaces of the drum, thus increasing total heat dissipation. The thermal conductivity of the constituents of cast iron in cal/cu cm/sec/C has been given as follows by Tophill & Angus:[4]
>
> Ferrite 0.17 Pearlite 0.12 Carbide 0.017 Graphite 0.33
>
> Since we have good reasons usually to avoid the presence of ferrite and carbide in our best brake drum irons, the conductivity of pearlite and graphite is one reason for keeping this constituent high in brake drums. (See SAE and ASTM 113, 114, and 115 irons.) Austin indicates that silicon reduces conductivity in iron.[5] This also is one reason for preferring low silicon.
>
> 9. High but stable coefficient of friction is desirable to minimize fading and give low line pressure with consequent low mechanical stress and distortion of drum due to brake shoe pressure. Also, this tends to reduce 'stick-slip' with its accompanying roughness and/or noise. Uniformity of chemistry, physicals, and microstructure in the iron are expected to favor stability in friction coefficient.
>
> 10. High specific heat to store a maximum of Btu with a minimum of temperature rise. Obviously, higher specific heat would help minimize heat checking, thermal stresses, and the possibility of braking surface-layer temperature rise which can cause phase changes in the iron and even slight melting. As far as we know, very little can be done metallurgically to modify this property in cast-iron drums.
>
> 11. Low coefficient of thermal expansion, to lessen thermal stresses. Hence again we know very little about how to change this property metallurgically.
>
> 12. Low modulus of elasticity, lessens stress concentration, especially that caused by sharp thermal gradients. Modulus decreases considerably with temperature rise, from about 12×10^6 at 25 C to about 9.4×10^6 at 500 C. Low modulus reduces the tendency to heat check by reducing the possibility of permanent set or plastic upset in the weaker portions of the layer of metal nearest the braking surface of the drum. Since it actually has been observed that such plastic upsetting is the first step in heat-check formation, we can and do make low modulus iron for application to brake drums where heat checking is a problem. This feature is specified into SAE and ASTM 113, 114, and 115 irons through the medium of a minimum total carbon and control over combined carbon, leaving a relatively high graphitic carbon that produces low modulus.
>
> 13. High melting or softening point, to avoid rapid wear and possible though rare seizure of brakes.
>
> 14. Adequately high strength over a wide temperature range, for obvious reasons. Low phosphorus, sulfur, and silicon plus the judicious use of carbide stabilizing alloys help.
>
> 15. Resistance to scoring, to get more even braking and better lining life as well as drum life.
>
> 16. Good wear resistance, to achieve longer service life.
>
> 17. Resistance to heat checking, since heat checks tend to become larger until they may or do cause failure by rupture entirely through the braking section of the drum.
>
> In recent years, the presently specified brake drum irons 113, 114, and 115, standard for SAE and ASTM, have been developed for problem applications, jointly by many of the metallurgical and brake engineers who are most familiar with the problems involved. It is believed that they represent a consensus of the best thinking known to date to solve (as they now exist) these brake drum material problems in a practical way."

Cast iron has long been known as the brake drum material which presented the best braking surface. It also has other advantages which are not so clearly understood. The recent work done with aluminum on passenger-car drums and to a lesser extent on truck drums would lead some people to believe that cast iron may soon go by the boards as a drum material except for the braking surface. This is far from the truth. Fig. 4 shows the comparative advantages of a typical aluminum alloy and a typical cast iron. Note that only on two counts does aluminum excel when properly compared to cast iron.

Cast iron has, in addition to the better braking surface:

1. Lower cost per lb and per cu in.
2. Higher heat storage capacity per cu in.
3. Much higher emissivity.
4. Much less expansion with heat.
5. A much higher melting point.

Aluminum excels only in heat conductivity and in heat storage per pound. This latter advantage is offset by the fact that 1 lb of aluminum occupies more space than 1 lb of iron, and the space available for brake drums is generally limited.

Testing

Before a prototype drum is released for production, it must be thoroughly tested for the various performance and thermal characteristics required in a truck or trailer brake drum, as well as for its ability to stand up in normal service. Such tests can be partly or completely conducted as road tests made to specified schedules on level roads or over hilly or mountainous terrain, but it is much more expedient to run them on a dynamometer.

Dynamometer tests are more capable of being duplicated because they are relatively unaffected by

---

[3] E. L. Morrison, The Budd Co. Private communication.
[4] "Cast Iron Brake Drums," by Tophill and Angus. *Automotive Engineer*, October, 1953.
[5] "Flow of Heat in Metals," by J. B. Austin. American Society of Metals, Cleveland, 1942.

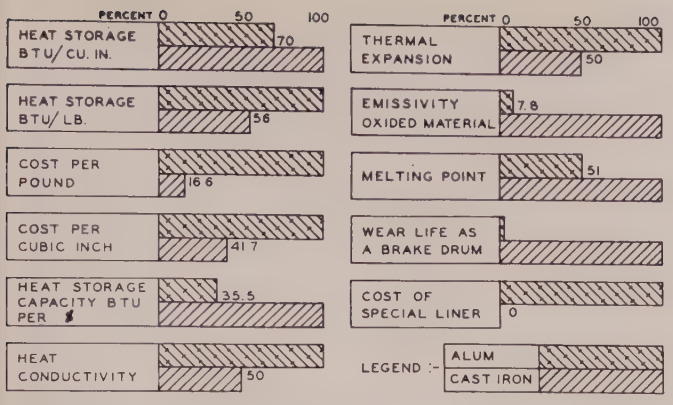


Fig. 4 — Brake drum material comparison, cast iron versus cast aluminum

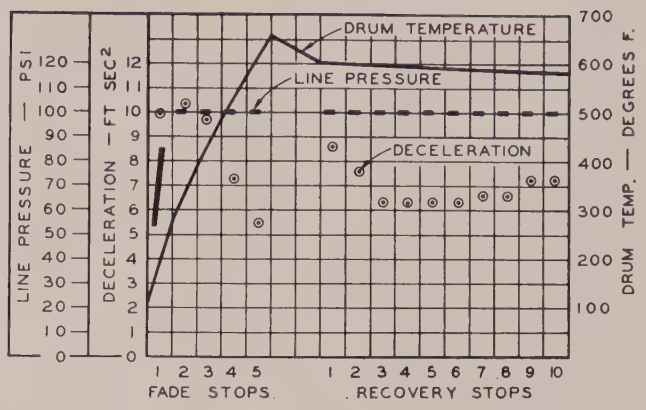


Fig. 5 — Typical 70 mph fade and recovery test

atmosphere conditions, whereas road tests are vulnerable to variations of weather such as temperature, wind, humidity, rain, or snow. Likewise, in repetitive stops the time interval of brake application can be more closely controlled; and since only one brake and drum is being tested, there is no question about the amount of work being done by it as there is when two, four, or six brakes are in operation on one test vehicle. Dynamometer tests are also faster since artificial air cooling, without the wheel shrouding the drum, is utilized to shorten the interval between stops in all but the thermal tests. In general, the preparations and some of the test procedures used by us are based on the SAE Recommended Practice — Dynamometer Brake Test Code. Following are some points thought worth noting.

Usually, a new drum is set up with the brake on the dynamometer with new lining blocks "Brake Doktored" to fit the drum. The drum has had three thermocouples, spaced at the centers of equal thirds across the width of the drum and 120 deg apart around the drum, welded in $\frac{3}{8}$ in. flat bottomed holes, drilled to 0.090 in. from the braking surface. The thermocouples are connected in parallel, to read average temperature, and the leads brought out through slip rings to a potentiometer circuit temperature recorder.

The dynamometer inertia load is set as close to the rated load as possible. A brake for an 18,000-lb axle is tested at 9000 lb.

The initial run is to burnish or wear-in the lining and consists of stops from 30 mph at an 11 fps^2 stopping rate. All stops are made after the drum has cooled to between 125 and 150 F. The burnishing continues until both shoes have a minimum of 80% contact with the drum.

The effectiveness tests consist of a series of stops at increments of 10 psi from 10 to 100 psi from 30, 50, and 70 mph. All stops are made with the drum temperature between 125 and 150 F; and the stopping time, stopping distance, and average torque are measured and used in computing deceleration. This test is made without a wheel mounted on the hub over the drum and with the maximum amount of cooling air circulating over the brake and drum. Brake noise, squeak, roughness, and "firebands," if they occur, are noted.

Fade develops in service either from continuous application of the brakes while descending long grades or from repeated stops or snubs on level road or on grades. This can be duplicated on the dynamometer by either continuous brake application under power or repeated snubs or stops with inertia load and no power. The complete stop method is used because it is more accurate and reproducible — since with a fixed inertia load, the prime controlling factor is the dynamometer speed when the brake is applied.

The fade tests are made by making 10 stops from 50 mph at 1-min intervals at 10 fps^2 deceleration. As fade sets in, the brake actuating air pressure is increased to hold the 10-ft rate, up to 100 psi maximum, which means with additional fade, the deceleration rate drops below 10 fps^2. After the tenth fade stop, the speed is reduced to 30 mph and 10 recovery stops, starting 3 min after the last fade stop, are made at 1-min intervals again holding 10 fps^2 deceleration or 100-psi maximum brake cylinder pressure.

After cooling the brake and drum to room temperature, a second fade and recovery test is made by making five stops from 70 mph, followed by a similar 3-min cooling and 10-stop recovery run.

For the fade and recovery tests, an inner wheel is mounted over the drum to produce the shrouding effect it has in shielding the air from the drum. An airstream is directed at an angle into the brake to simulate the airstream passing the wheel and brake on the vehicle. In service the airstream passing the wheel becomes turbulent when it strikes obstructions such as the axle, cam shaft, air cylinder, and their brackets, springs, and hangers. Checks of the dynamometer tests against road tests have shown this system of directing cooling air against the brake and drum to be satisfactory because of their close agreement.

Fig. 5 is one of a typical fade and recovery test on a 16½ in × 7 in. brake.

The steady state temperature — power input is also conducted with a wheel over the drum and a 30-mph airstream directed against the brake and drum. It consists of a series of stops from 50 mph at 10 fps^2 at 6-min intervals until the "valley" or starting temperature reaches equilibrium. Similar series of stops are made at 4-, 3-min, and shorter intervals. Steady state or equilibrium temperature is attained most quickly by making three or four stops at the beginning of each series to bring the brake and drum to a temperature above its eventual

Table 2 — Calculated Continuous Brake Horsepower

Intervals, min	Hp	Intervals, min	Hp
6	3.80		
4	5.70	1¾	13.05
3	7.60	1½	15.20
2½	9.10	1¼	18.25
2	11.40	1	22.80

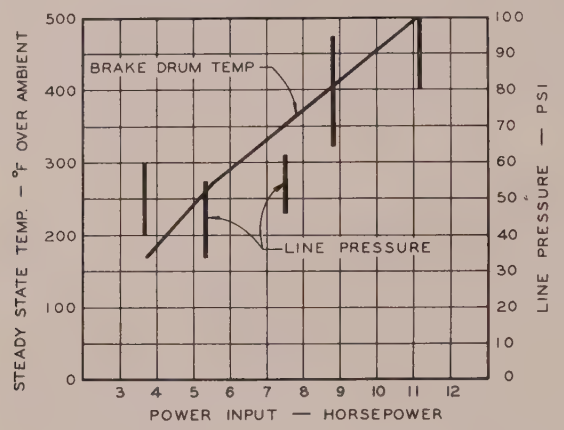


Fig. 6 — Typical steady state temperature — power input test

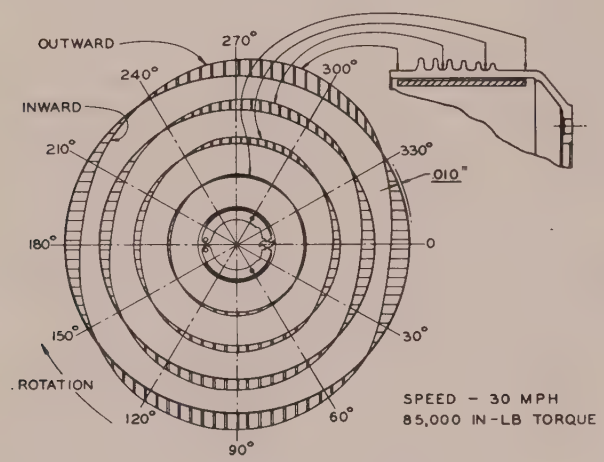


Fig. 7 — Brake drum distortion

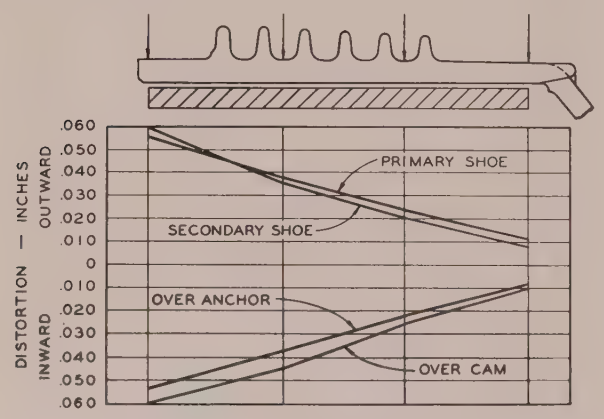


Fig. 8 — Brake drum distortion (maximum)

equilibrium temperature.

Table 2 shows the calculated continuous brake horsepower for a 9000-lb load braked from 50 mph to a complete stop at the standard test intervals.

It has been noted that this steady state temperature — power input test is a measure of the heat dissipating ability of the brake and drum and is independent of thermal capacity of the brake and drum. The test also shows the capability of the brake and drum to operate at elevated temperatures for long periods of time.

Results of a typical steady state temperature power input test are shown in Fig. 6 for the 16½ in. × 7 in. brake and drum.

The physical distortion of a brake drum during application of the brake is probably one of the main factors in determining whether the drum is properly designed for the particular brake in question. Unfortunately, the effects of drum distortion on performance and life are not fully known. This distortion is only that resulting from the drum being pushed outward as it passes over the shoes and being pulled inward between the shoes. Naturally, it is greatest at the open edge of the drum and smallest at the drum head side. Distortion, as considered here, excludes thermal expansion and its allied stresses and strains due to the temperature gradient throughout a drum. The gradient at times may be quite large since the heat of friction is developed on the braking surface and conducted to the extremities of the drum, such as outer edge of ribs or drum head, with appreciable delay.

It is believed that insufficient drum distortion results in low effectiveness, probably because of unequal distribution of pressure of the drum over the lining blocks. It is also conceivable that too rigid a drum might be rough or noisy. Conversely, too much distortion allows the air cylinder to bottom at high pressures, resulting in insufficient shoe pressure and effectiveness. It is also believed that a flexible drum causes uneven lining wear, but this is most likely averaged out by the high and low rate applications of the brake.

Drum distortion is measured by a displacement indicating transducer such as a lineal variable differential transformer or strain gages on a cantilever beam. The transducer is mounted on a solid bracket attached to a wheel disc mounted on the hub flange. The transducer stylus rests on spots filed smooth on the outer surface of the drum. Leads from the transducer are brought out through slip rings and fed through an amplifier and into a Brush recorder.

Typical drum distortion charts are shown in Figs. 7 and 8 for a 16½ in. × 7 in. drum during a 30 mph, 15 fps² stop.

Establishing a reliable brake drum life or durability test is a real problem because of the many contributing factors that eventually cause brake drums to be replaced. Probably the most important of these is drum wear and the closely allied scoring. Since wear is mainly due to abrasives getting inside the brake, a test to duplicate these wear conditions would be difficult to set up, time consuming, and its value in producing better brake drums is questionable. Some drum wear is also attributable to abrasive materials in the lining.

Heat checks are the next most important reason for replacing brake drums; first, because they increase lining wear, and second and more important,

because they lead eventually to cracked or broken drums. Heat checks develop at high-temperature differentials between the braking surface and the surrounding material and are easily produced on the dynamometer during high-speed, high-rate stops.

Cracks may also develop due to overstress. In new drums this could be caused by poor or weak cast iron, imperfections, or improper design. In old drums it might be attributable to drums worn or rebored too thin, to cooling after the brakes had been set while hot, or to heat checks that had progressed too far. These cracks are produced on the dynamometer under the same condition of high pressure necessary for the high-rate stops of heat check testing. Because of this, it is difficult to set up an accelerated fatigue or durability test for brake drums.

Strain gages are of some help in locating high stress points in brake drums, but their use in measuring dynamic stress is limited by the temperature limitations of the strain gages and by the fact that too often the maximum stresses in a drum are on the braking surface.

The ultimate worth of a brake drum is never evaluated until hundreds of them have been operated by various fleets, over all parts of the country, with many kinds of brake lining with various degrees of maintenance and in all kinds of weather.

Brake Drum Costs

No paper of this type can avoid the dollar sign on brake drums. Any successful automotive product must do a good job at the lowest possible price. Competition enforces this, but customer pressure is apt to cause some relaxation at times in fundamental principles. The desire on the part of the operators of commercial vehicles for ever increasing pay load and continual reduction in tare weight poses a real problem. Weight reduction in brake drums without increased cost or reduced performance is the goal of all. That weight reduction, without other compensating changes is dangerous and costly can be verified by those who keep records. As an example, one company, according to Mr. Bassett, ". . . tried it this year on 60 tractors, and had to replace over 200 broken drums in less than 8 months. They corrected this breakage by first going to 7 in. wide drums on the 6 in. wide brakes, then to a heavy 6 in. wide drum that is 30 lb heavier than the ones causing trouble. Since the change to heavier drums, the heat checking, which cut the friction material life in half, and drum breakage have been eliminated."[6]

The corrective action taken by a safe estimate required the expenditure of $150 per vehicle for material plus at least $50 for labor for a total of $200 without counting vehicle down time. We feel certain that their own records would show a much higher loss.

A recent survey on brake problems and costs conducted by the "Fleet Owner" shows conclusively that too many fleet operators keep too few important records. Good records will help to pave the way to better braking. Those interested should take advantage of the help available through trade organizations such as the T.T.M.A. or the A.T.A., or through other sources such as the Commercial Car Journal.

Reboring of Brake Drums

Brake drums are designed with a factor of safety for heat storage capacity and physical strength. However, pressure for reduced weight on the part of the ultimate user has caused a reduction in this factor of safety in many instances. This means that drum re-conditioning programs should be reviewed in this light. Reboring of brake drums to remove scoring and heat checks can be carried too far, in fact, to such an extent that the heat storage capacity is reduced to dangerous levels. Obviously, when wall thickness is reduced, the ultimate strength of the drum is also reduced. This combination of reduced heat storage capacity, which promotes higher brake drum temperature and, therefore, brake fade, and reduced physical strength can cause premature drum failure. This is particularly true where heat checking is involved.

Our recommendations on re-conditioning medium-sized truck and trailer brake drums would prescribe as a maximum, an increase in diameter at rebore of ⅛ in. This permits two rebores of 0.030 in. per side or a total diameter increase of 0.120 in. Lightweight, heavy-duty commercial drums should be limited to the same amount since, in the interest of decreasing weight, they have but little inherent safety factors.

Reboring heavy truck and trailer drums should be judged on the basis of reduction in drum ring thickness. We recommend the drum ring never be bored to reduce it below seven-eighths of its original minimum thickness, and this only in cases where the original thickness is ½ in. or greater.

Excessive reboring can be a short-sighted policy aside from the safety angle, in that brake-block damage is usually incurred when drums break. This is a readily measurable monetary loss.

Conclusions

1. An engineering term to define brake, brake lining, and brake drum heat dissipating ability is necessary. We hope that we have helped in this respect with chp.

2. Cast iron is an excellent brake drum material. We believe that it will be used for some time to come.

3. Brake lining material is our present weak link. Unless a major breakthrough in this area is forthcoming, no major strides in reducing both the weight and cost of brake and drum assemblies appear possible.

4. Under the limitations imposed by current materials, the major element not yet sufficiently appreciated by the industry is drum positioning and exposure.

5. Dynamometer testing of brakes and brake drums is more precise than highway testing, but cannot be used exclusively. The final testing of all elements involved must come from actual use on the highway.

---

[6] "Motor Vehicles Maintenance Course," by J. V. Bassett. Purdue University, October, 1958.

Discussion of this paper will be found on p. 349, following the Super paper.

economics of Heavy-Duty Brake Design and Maintenance

R. K. Super

Rockwell Standard Corp.

THE TITLE of this paper infers an inter-relationship of design and maintenance in the overall economics of vehicle brake equipment. The need is paramount for a more comprehensive understanding of the design factors and maintenance problems involved by both the engineering and maintenance groups if the full economics of the current brake designs are to be realized. Any attempt to justify marginal engineering by claiming poor maintenance is as biased as the opposite point of view. Both positions fail to solve the basic problems, but do focus attention on a requirement for improvement in both design features and maintenance practices.

The ultimate in foundation brake design is ideally defined by stating that the brakes should be adequate and last the life of the vehicle with no service requirement. This concept is somewhat visionary at this time; but the idea of a brake that requires no attention between lining changes is already an established reality in passenger-car brakes and is in the development stage for heavy-duty commercial vehicles.

The elements of such a brake design are reasonably simple and would be as follows:

1. Maximum lining life.
2. No periodic lubrication.
3. Weather sealed for uniform performance.
4. Automatic adjustment.

It is reasonable to feel that extended lining life is the important single aspect of the economics of foundation brake equipment, but the broader concept is outlined in the following abstract from a report by the AMA-TTMA Brake Committee dated March 9, 1955:

> "Agreement was reached on the following points relating to an elaboration of that phase of the problem regarding extended brake service life and periods between brake adjustments:
>
> 1. Maximum capacity in the brakes on the vehicle or combination may be obtained by providing brakes on all wheels.
> 2. Maximum capacity of brakes on the vehicle or combination requires synchronized braking on all axles.
> 3. The use of the largest size brakes within the space limitations should be a further objective.
> 4. Maximum cooling of the brakes through consideration of wheel design, tire-rim sizing, and wheel shrouding would be a further contribution to extended service life.
> 5. Power-actuated components should be selected to provide maximum reserve and be limited by the space available and the possible penalizing of the application time.
> 6. Consideration of the use of supplementary brakes such as retarders, to obtain extended service life for the vehicle brakes, should be a continuing study.
> 7. The introduction of limiting downhill speeds to handle the vehicle or combination within the capacity of the brakes would be an important contributing factor to extended service life.
> 8. A realization that more complex problems involving the economics of brakes have resulted from higher scheduled speeds, greater engine horsepower, and higher gross vehicle and combination weights."

Design and Operational Problems

An elaboration of these points in the sequence listed will indicate further some of the design, operational, and maintenance problems:

1. The use of brakes on all wheels has been recognized as desirable, but not completely acceptable, because of the problem of front brake control. The proportioning of front brake power to secure optimum brake control and capacity, free of driver selection, would add immeasurably to an improvement in vehicle braking through the efficient use of brakes on front wheels.

2. The synchronizing of brake power, as applied to the wheel units, involves two phases. The first relates to the delivery of the actuating medium at the same time at all locations in the system. Suitable valves and installation techniques are now available which make possible the practical realization of this objective. The second phase of synchronizing is the problem of power distribution. This is particularly critical on brakes for commercial vehicles because of the extremely variable conditions of loading. In addition to the difference in vehicle

THIS PAPER presents a review of the design and operational problems of heavy-duty truck brakes. One of the major development goals is brakes that require no attention between relinings (as are now available on passenger cars).

The author discusses point by point the AMA-TTMA Brake Committee agreement relating to extended brake service life and periods between brake adjustments.

Emphasis is placed upon maintenance programs which provide for frequent inspection of the vehicle. The margin of brake performance deterioration is narrow.*

load distribution between a fully laden condition and a completely unloaded unit, the partially loaded or unequally loaded vehicle or combination is even a more serious problem. Much successful work has been done on "load valves" which proportion the brake power in relation to the load on the wheel.

3. The use of the largest brake possible in the space available has been the engineering basis for obtaining greatest service life. This does not necessarily mean maximum brake lining area, as this factor does not directly indicate brake capability. The brake size also is related to the efficiency of the design when operating in the space available. On many installations a brake of smaller diameter will have greater life than one of larger diameter because of tire-rim size. A combination of brake size and the design of the companion brake drum will produce the best results.

4. Added brake life is possible through careful attention to the environment in which the brake operates. Natural air circulation both inside the brake and over the brake drum should be made possible by exposure to the airstream. This means minimum use of dust shields on the brakes or wheel shrouds which prevent air circulation.

5. The selection of power units to actuate the brakes can contribute much to providing added reserve travel in the actuating components to provide a greater margin for "follow-up" which is necessary to compensate for drum expansion and secure uniform performance with less frequent adjustments for lining-drum clearances. Overdesign wherein power units are used which are needlessly large will reflect in a greater stopping distance. It is in this phase of brake design that automatic clearance adjustment is very much needed. Recent progress on this item has been very encouraging.

6. The use of "retarders" as a means of reducing service brake costs is complicated by its cost, weight, and space required for installation. The adverse load effect on drive line components has also had a deterring influence regarding their wider use. The need for this type of device is explained by the desire to negotiate grades at speeds in excess of those which would permit the use of the engine in the lower transmission gear ratios where the engine is effective as a load retarder. It is the loss of engine "holding" capability at desired road speeds that is a major contributing factor to present service brake problems.

7. The optimum capacity of all the braking units on a vehicle would be that which permitted the safe control and operation of the vehicle on any grade at any desired speeds. The vehicle load and speed in combination with the grade would establish the required brake capacity. Recognizing these basic factors, it is understandable that safe and economical vehicle operation requires that the loads and speeds be controlled within the limitations of the ability of the available brake equipment. Legal limitations on both factors do not necessarily make the brakes completely adequate for these "allowed" conditions. Further, it is axiomatic that any vehicle brake which will safely stop the unit is adequate in capacity for some specific vehicle load and speed. These may be lower than desired, in which case the economics of safe operation will establish the relationship of brake size or capacity to the vehicle speed and weight.

8. The economics of brake equipment have been seriously complicated in the past few years by higher horsepower engines and automatic transmissions. By virtue of this combination, it is possible to accelerate more rapidly and the time intervals between brake applications are reduced. It is also possible to ascend grades at a higher speed with subsequent reductions in time. Actually, the "uphill" time to climb a grade is the "cooling-off" time for the brakes; and as this time is shortened, the temperature of the brakes is higher at the start of the next descent.

Maintenance Practices

This review of design and operational problems as they relate to the economics of vehicle brakes can now be supplemented by certain maintenance practices which can influence the ultimate results.

The first problem of the maintenance group is to service the vehicle brakes so that they meet the performance requirements of the various State and Federal regulatory or inspection agencies. In this respect, the brakes must be maintained between the original new performance level and the level required for either *legal* or *safe* traffic performance. The engineering concept of commercial vehicle brake power which is the original or new condition of the brakes is predicated on two factors. The first is the coefficient of adhesion of the tires and the road. Based on published information by the automotive tire manufacturers, it is established that a 0.6 adhesive factor constitutes an average figure representative of partially worn tires in their relationship to an average road surface.

The second factor relates to the safe handling of the load on the vehicle. In the case of commercial vehicles, the body design and characteristics of the load largely determine the maximum permissible rate of braking. In many instances, the rate of braking must be limited so that the load will not slide along the floor and overload the vertical body

* Paper presented at SAE National Transportation Meeting, Chicago, Oct. 27, 1959.

members, such as the headboard, to the point of failure. Strapped or chained loads also must receive similar consideration. Through many years of actual usage, it is quite universally established that a 0.6 adhesion factor between the tire and road is practical. This 0.6 factor corresponds to a rate of deceleration of 19.3 fps^2. If it was possible to maintain this rate uniformly throughout the stop the measured distance would be 22 ft approximately at a speed of 20 mph. This figure has found wide usage in brake engineering for commercial vehicles and relates to the brake power originally built into the brakes. This may be termed the full or 100% performance of the brakes.

It cannot be expected that a vehicle braking system can be maintained at its full original efficiency and performance. It is, therefore, important to recognize the permissible extent of deterioration. The minimum acceptable performance of the brakes must be based on what constitutes safe brakes. Law enforcement groups have indicated this requirement, but actually that minimum braking ability which will adequately meet traffic requirements and provide a margin for emergency conditions such that safe and economical operation of the vehicle is secured is the desirable objective.

Traffic surveys have indicated that a deceleration rate of 8 fps^2 covers the average braking rate for various types of vehicles, that is, busses, trucks, taxicabs, streetcars, and passenger cars, when operating in either congested traffic or boulevard and open highway driving. These surveys also reveal that emergency or panic traffic stops seldom exceed 12–13 fps^2.

Supported also by extensive surveys and actual tests, the Interstate Commerce Commission require carriers under its jurisdiction to maintain brake performance at the level outlined in the Uniform Vehicle Code of a decelerating rate of 14 fps^2 or a measured stop of 30, 40 or 50 ft depending on the type of vehicle.

The brake design capability of 19.3 fps^2, as compared with a required minimum rate of 14 fps^2, indicates the margin of performance deterioration of the brakes. The brakes, therefore, must be maintained at or above 75% of the original efficiency. This is the margin for maintenance.

This is a relatively narrow range in which to maintain equipment. The first impulse is to increase the performance at the upper limit but, as previously mentioned, this introduces load control problems — and probably equally important, the control of the unladen vehicle. This limited range of brake performance at 75–100% efficiency becomes the basic maintenance requirement and the responsibility of the maintenance group. The actual cost of this maintenance may very well be the vehicle brake cost per mile of operation. Quite conceivably, it could be supplemented by accident costs resulting from faulty brake equipment.

The maintenance department should be given vehicles of a type best suited for the particular operation. Undersized and underbraked vehicles are a continuous expense. The following points might suggest the features of greatest interest to the truck maintenance man:

1. The vehicle should be equipped with brakes having the potential of a long service life between relines and with a minimum requirement for lubrication and adjustment.

2. The brakes on all units should be synchronized in regard to airflow timing and power balanced in relation to the loads.

3. The environment in which the brake operates should insure the lowest possible brake operating temperatures.

4. The loads, speeds, and grades to be negotiated must be within the economic capabilities of the wheel brakes or a supplementary brake (retarder) will be a necessity for economical operation.

The first item is a "built-in" feature of the vehicle brake equipment. The vehicle either is equipped with adequate brakes or the servicing of these parts becomes a major expense to provide the necessary performance.

The second item relating to synchronization of brake actuation and power balance is largely a maintenance problem, assuming that the original equipment was properly engineered. The problem of maintaining good brake timing is one of keeping the control valves in good condition by periodic checks and overhauls. This guards against corrosion and leakage. Brake lines, flexible hoses, couplings, and the like, require similar attention. The advantage of synchronized brakes is the extended service life obtainable through the distribution of the total braking load to all of the brakes on a vehicle or combination. The introduction of a hand control valve for independent operation of the trailer brakes completely nullifies the concept of synchronized or balanced braking and contributes immediately to a high maintenance requirement on all brakes on the combination.

Balanced braking is the application of power to the wheel brakes to produce a uniform retarding rate to the braked wheels. This feature is engineered into the vehicle but requires continuous maintenance if it is to be retained. The initial contributing factor to unbalance is the failure to maintain uniform lining-drum clearance on all of the brakes. This varies the stroke of the power units to upset the power balance. It also changes the volume requirements to actuate the brakes, thus upsetting the timing. Frequent and periodic adjustment of all brakes is good maintenance.

Following brake adjustment in importance is the use of brake linings for replacement which are either the original equipment material or the "equivalent." The use of friction materials which have not been thoroughly evaluated can lead to serious unbalance. Even when proper brake linings are installed, the best results will be obtained only if all of the brakes are relined or resurfaced at the same time.

In addition to maintaining brake balance by proper adjustment and relining practices, a similar consideration is necessary with regard to brake drums. The weight and design of the brake drums are equally important in providing good brake balance. Replacement drums should be comparable in these respects and preferably changed in pairs at both ends of the axle. This is applicable to the use of turned drums also.

Item three relating to brake environment brings out one point in brake maintenance on brake designs utilizing detachable dust shields. In general, these are removed in summer weather and installed in the winter. If lower maintenance costs resulting

from better lining mileage is obtainable with either option, then it should be adopted.

Item four indicates that vehicle loads and road speeds and grades are of concern to the maintenance group. Actually, brake maintenance costs are directly related to these factors of operation. It could, therefore, be concluded that the maintenance department would be in the best position to make record-supported recommendations regarding the economies of vehicle operating costs in their relationship to these operating factors.

Conclusion

It is evident that brake equipment for heavy-duty commercial vehicles must be initially designed and engineered to incorporate those elements of size, or capacity, and brake balance which are conducive to low maintenance costs. The operating conditions of the vehicle or combination must be within the capability of the selected brake equipment to make economical maintenance possible. Finally, the maintenance programs and techniques must recognize the narrow margin of brake performance deterioration which makes necessary frequent periodic inspection procedures to insure the safety of the vehicle. Full evaluation of the original engineering choice of brake components and materials should not be discredited or changed without an equivalent background of supporting information.

As in most undertakings, the complete cooperation of all interested parties achieves the best results. In the transportation business, the joint efforts of the design engineer, the truckdriver, and the mechanic can result in the greatest economies of vehicle operation. New ideas will make the job of each somewhat less exacting but the available equipment is the present problem and that of the future will present new, more complex problems.

DISCUSSION OF HYKES-HERMAN AND SUPER PAPERS

Maintenance Man States His Side of Brake Problem

— E. A. Bruha
Gateway Transportation Co.

THE VIEWS of any maintenance man are necessarily conditioned by the type of operation in which he is engaged. Since we operate primarily over the highway in the relatively flat terrain of the Middle West, our requirements and our general feeling toward brakes are quite different from someone operating off-the-highway or in mountainous terrain.

Early in his paper Mr. Super lists the elements of the ultimate in brake design as maximum lining life, no periodic lubrication, weather sealed for maximum performance, and automatic adjustment. He goes on to say, "Extended lining life is the important single aspect of economics of foundation brake equipment."

I agree that getting maximum lining life is important and, as far as the actual foundation brake itself is concerned, this may be the most important single item. It is, however, a relatively small part of the total cost of maintaining vehicle brakes. In a fleet such as ours, the biggest expense and the most maintenance problems arise out of the multitude of valves and piping that make up the brake actuation system. Mr. Super indicates that major increases in lining life have to come through better synchronization of vehicle brakes and equipment to adjust the braking effort with load variations. If this can be done through improvements in the foundation brake, I am all for it. If it means more valves, I think we will end up spending more to gain these two ends than will be realized in lining savings.

Synchronization and equalization of brakes have, of course, other advantages than improved lining life, including better vehicle control and improved tire life. Unfortunately, vehicle manufacturers do not yet take this seriously. They take basically good components and then do all they can to sabotage their operation with poor piping.

I agree wholeheartedly with Mr. Super's comments concerning improving lining life through better cooling of the brake systems, better lining materials, and better drums.

In reviewing Messrs. Hykes and Herman's paper, I was very interested in the work being done to establish a standard rating for brake drums. I only hope that this can be developed to a point where all brake drums will be rated and a purchaser will know what he can expect from a drum at the time he buys it. Brake drums are very much alike in appearance and it is next to impossible to tell a good drum from a bad drum merely by looking at it. About the only way he can tell which one is good or bad is by extensive field tests. Then, by the time the test results are in, all of the manufacturers have new and supposedly much improved drums available and the test results can no longer be applied.

Wear is, of course, the biggest reason for brake drum replacements. We all recognize this as a necessary evil that occurs when you rub two materials together. The authors' remark concerning the effects of temperature on wear makes a good case for better cooling of the drums.

Other than wear, our biggest problem is with heat checking and cracking, which account for at least 35% of our drum replacements. I am inclined to think that this is as much a driver problem as it is one of maintenance or drum design. We always hope that the drivers will use their brakes properly but, since that is not always the case, we would like to have a drum design that takes into account driver abuse as well as normal wear and tear.

Messrs. Hykes and Herman mention that the continuous application of brakes in descending a long uniform grade is not desirable for highway service. I have always felt this somewhat more desirable than a series of brake applications, if for no other reason than that the amount of air available to a driver is somewhat limited. It is also my belief that a short hard application was much more likely to set up drum surface temperatures which would result in heat checks than would a longer but less heavy application.

Rating and Maintenance Of Brakes a Continuing Problem

— R. A. Goepfrich
Bendix Products Division, Bendix Aviation Corp.

THE AUTOMOTIVE INDUSTRY has had a long-standing need for a good, simple method of rating brakes. The longer we study this problem, the more convinced we become that there may not be a real simple method. This is evidenced by the fact that the SAE has had a Brake Subcommittee working on this problem for at least 12 years; this committee is still active. An early solution is urgently needed to stymie the type of legislation mentioned by Messrs. Hykes and Herman and the tests which the AMA-TTMA Brake Committee are conducting is a big step in this direction.

This proposed method of rating brakes is derived by running tests on the vehicle, because operating conditions cannot be fully duplicated on the dynamometer. For ex-

ample, we know of no dynamometer setup in which a whole set of vehicle brakes can be tested simultaneously. Nor one in which cooling conditions can be exactly duplicated to simulate the environment of the brake in the vehicle. It is hoped that eventually, when a considerable amount of data has been accumulated from vehicle and dynamometer tests, it may be possible to establish factors which can be applied to dynamometer results to correlate them with the vehicle.

As the authors point out, neither radiation nor conduction gets rid of very much heat; nor can we expect very much improvement in heat dissipation due to different drum irons. Thus, the best chance for improvement is to provide for better convection. This poses somewhat of a dilemma for the brake engineer because on the one hand he should provide as much air space around the drum as possible, and on the other, he wants to get in a brake as large as the wheel will permit (as Mr. Super pointed out). Consequently, he must choose the best compromise.

Since the heat dissipating system is made of two main elements, the lining and the drum, they must both be considered in the overall performance. In the lining, the coefficient of friction determines its effectiveness and the composition determines its life. In the drum, the material of the rubbing surface determines its cofficient, and the operating temperature of the brake determines its efficiency as a heat dissipator. Small improvements in each of these elements, although insignificant in themselves, may show a considerable improvement when put together.

For example, on linings, there is considerable work being done on: (1) improving organic linings, (2) combinations of ceramics and metals, and (3) combinations of cerametalic and organic linings.

In drums, improvements can be realized by:

1. Increasing heat dissipating area by means of ribs or flares, or both.
2. Positioning in the airstream as well as possible.
3. Artificial cooling by air, oil, or water.
4. Use of materials with high heat transfer coefficients.

As was previously mentioned, the problem of providing a brake system with adequate performance, satisfactory life, and ease of maintenance may not be solved by any one of the above-mentioned items. However, consistent improvement in all the areas, plus skillful tailoring of one to the other, should result in economies which will allow this basic brake system to live for some time.

Maintenance

If the design objectives which Mr. Super so clearly described were fully met, it would go a long way toward easier and better maintenance. However, since this ideal situation has not yet been achieved, we must strive for "a more comprehensive understanding of the design factors and maintenance problems by both the engineering and maintenance groups," as pointed out by the author.

This means that engineers should spend some time in the shop with the maintenance men and learn first-hand what their problems are. This knowledge should then be used to improve the design for easier maintenance.

On the other hand, the maintenance men should strive to become thoroughly familiar with the design and maintenance requirements, especially when a new brake is involved. Too often, previous experience and intuition alone are relied upon to determine maintenance procedures. Another prevalent practice, and a bad one, is that technical literature and service manuals get buried in the shop foreman's file instead of being in the hands of the men who actually do the work.

The ideal braking system is one in which each brake does work in proportion to the load carried by its wheel. The big differences in load distribution between the laden and unladen vehicle precludes perfect braking distribution under all conditions of loading. Hence, a brake must develop enough power to safely handle the fully loaded vehicle and yet not so much that the brake locks up at comparatively low decelerations in the unladen vehicle. We hope that a good load proportioning valve is not too far off.

We should not overlook the fact that bad maintenance practices, such as neglecting the brake adjustment on one axle, can greatly upset the braking balance.

Since extended lining life is one of the important aspects of the economics of brake maintenance, we should recognize the effect of elevated temperatures on it. We know that when the critical temperature of a lining is exceeded, it breaks down and wears very fast. Also, most linings "fade" at high temperatures because their coefficient of friction decreases. Under these conditions, higher input pressures are required and this also produces more wear.

Therefore, the designer must provide a brake system which does not produce these high temperatures under reasonable driving conditions and the operator of the vehicle must avoid them by judicious driving habits.

Regarding cooling of the brakes to increase lining life, some recent tests indicated that positioning the brake drum so that its open end projects out from under the rim has shown some definite gains in this respect. This is a point which the designer should keep in mind.

During the last war the Ordnance Department had some extremely difficult service and maintenance problems on vehicles which operated in the South Pacific area. The combination of salt water, mud, and bad terrain shortened brake life to some fantastically low figures. Often the absence of nearby supply depots prevented servicing and maintenance entirely. For this reason, development of a completely sealed brake is under way.

One design which we are familiar with not only completely seals the brake to prevent the ingress of dirt, mud, and water, but has the brake drum on the outside of the wheel. In addition to affording excellent cooling, the brakes and drums can be removed without jacking up the vehicle. This makes replacement of linings comparatively simple.

Although this design requires some rather extensive vehicle changes, if it achieves its objectives of longer brake life and reduced maintenance, it might eventually become popular in the civilian field.

Basis of Brake Rating Formula Should Be Intermittent Use

— S. G. Tilden
The Permafuse Corp.

THERE IS one specific point brought out by Messrs. Hykes and Herman and others on brake effectiveness rating that has me concerned.

The method of rating present-day brakes by the maximum amount of energy that they can absorb and yet maintain constant temperatures, must result in comparatively low horsepower ratings, because the brakes themselves are not designed to operate at constant temperature.

The one particular 16½ in. × 7 in. brake described is shown as being capable of an instantaneous horsepower rating in the magnitude of 730 hp. And such a rating is adequate to bring the specific vehicle described to a stop within the Uniform Vehicle Code requirement. But the point that the authors bring out later (as have others before them) is that these same brakes cannot absorb such horsepower in continuous operation.

"Rated as a continuous operating device wherein an acceptable "steady state" temperature is maintained, the horsepower absorption capacity of this brake drops from the instantaneous 730 hp developed in Eq. 3 to a continuous 22.8 hp, as shown in Table 2.

Obviously, this is because there has been no cooling system designed into the components of this brake system to carry away the heat generated in continuous horsepower absorption. There could be, and are, such cooling systems designed into a brake to do this, — to wit; the water-cooled brake now under development by Wagner Electric Corp., and various retarder devices with built-in cooling systems

also under development. But these systems are expensive, complicated, and heavy.

However, any rating of the orthodox brake as a continuous duty device must necessarily be low, as pointed out previously. My contention is that such a method of brake rating would be too low for acceptance by State Legislators and Motor Vehicle Commissioners purely on the basis that 22.8-hp brakes are not adequate for a vehicle with a several hundred horsepower engine.

Then too, a continuous horsepower absorption concept must be recognized as being just what it is, one of many tests to be conducted on a brake system, bearing in mind that a retarder can have a very high continuous horsepower absorption rating, but *will not* bring the vehicle to a complete stop.

Since the orthodox brake is designed and used as an intermittent energy absorption device, I urge that any horsepower rating formula developed and promulgated be based upon the ability of the brake to absorb energy in intermittent use— not in continuous service use.

This problem of brake rating is a big one, and it is not new. SAE Brake Subcommittee IV has been working on a method of rating brakes for many years (my notes go back to 1947) and have under consideration the work done by an AMA-TTMA Committee in determining continuous horsepower absorption ability of various truck brakes at constant brake temperature. These tests bear out the contention that rating brakes as continuous horsepower absorption devices results in an unacceptably low figure and is not a measure of their true power.

A further bibliography on the subject would certainly include these papers:

1. "Brake Ratings for Automobile Vehicles," by R. L. Wehe. SAE Transactions, Vol. 65, 1957, pp. 219–231.

2. "Postwar Brakes," by J. G. Oetzel. Paper presented at SAE Metropolitan Section, New York, May, 1945.

3. "Brakes — and Brake Lining Characteristics," by J. G. Oetzel. SAE *Quarterly* Transactions, Vol. 4, April, 1950, pp. 217–231.

4. "Brake Designs and Methods of Rating Brakes for Commercial Automotive Vehicles," by R. K. Super. SAE Transactions, Vol. 54, 1946, pp. 205–214.

ORAL DISCUSSION

Reported by E. J. Keenoy, Jr.
Raybestos-Manhattan, Inc.

J. W. Riesing, Pacific Inter-Mountain Express: We have requested that manufacturers furnish truck operators with specifications for use in ordering replacement materials so that their equipment can continually meet the conditions imposed. On various occasions we have attempted to purchase the original equipment parts for replacement purposes with no success. We feel there is a need for the identification of parts which original equipment manufacturers recommend so that we may maintain the performance obtained from new vehicles.

We are looking for new designs that will be simpler. We are looking for a reduction in the number and complexity of air lines, valves, and the like. We will welcome any self-adjusting brake device which is practical.

Our brake block life since 1950 has been around 75,000 miles.

With respect to using original equipment specifications I would like Mr. Hykes' opinion on whether any difference in brake block and brake drum life can be expected if the fleet operator specifies original equipment materials for replacement.

With reference to Mr. Super's remark that engineers should spend more time in the field we welcome any engineering help that is available.

We are in the midst of a campaign to reduce weight. What is being done to reduce the weight of brake drums, brake components, and brake system accessories?

Willard Lord, Penn State University: Some operators have driver training programs pointed toward equipment conservation. In the training programs with which we have had contact, we have used as authority George Oetzel's work, and have advocated the use of a steady drag rather than intermittent applications.

Burt Ogden, Consolidated Freightways: We have advocated steady applications to our drivers to negotiate long grades.

Mr. Super: Several years ago we conducted tests in Jennerstown to determine what temperature difference could be attributed to the two methods of negotiating a long grade. These tests indicated about 25 deg lower temperatures with steady applications but it was felt that this 25 deg was within the normal measurement error and the tests were deemed inconclusive. We feel that this problem could better be evaluated by using the methods prescribed on several trucks over a long period of time.

William Birge: Should aluminum be considered as a brake drum material in view of its light weight and high heat conductivity properties?

Mr. Hykes: When we consider strength-to-weight ratio in addition to weight and conductivity, we feel that cast-iron drums will be with us for some time to come.

Cy Burton, Aluminum Co. of America: Aluminum's advantage as a brake drum is demonstrated by its being offered on both the Buick passenger car and the Pontiac pursuit car to reduce brake temperatures. Reports from General Motors indicate that brake lining life has increased 3–5 times because of the lower temperatures afforded by aluminum drums. Some time ago, Alcoa, developed a $16\frac{1}{2} \times 7$ in. aluminum drum with bolted in cast-iron liners. Consoliated Freightways ran these drums for 10 years satisfactorily. The cast-iron liners were replaceable. Aluminum with high silicon content is being developed for brake drums, but is presently not strong enough for truck use. There is also work being done on a forged aluminum drum with welded wear surfaces.

Dan Walther, Dayton Steel Foundries: (1) Given an average brake drum, can Mr. Hykes from his formula predict brake drum life? (2) Have you established a relationship between drum mass and life? (3) Does the depth of heat checks reduce drum life regardless of how much has been cut out of the drum?

Mr. Hykes: (1) Weight and horse power ability would affect drum life but ultimate life would depend on drum usage. (2) Mass is inconsequential in establishing drum life but is important during a fade when mass acts as a heat reservoir. (3) Boring out a brake drum decreases mass and would have the effects described above.

William Walther, Dayton Steel Foundries: With reference to Mr. Super's remarks about replacing components with original equipment materials, this is sometimes impracticable because of mixed vehicles in a fleet and the problems of matching tractors and trailers.

Suspension plays a big part in brake balancing. This is particularly true in cases we have noted where the front axle of a tandem was hopping and could not provide the same tractive effort as the tandem. We have noted that the rear axle of a tandem heat checks and wears faster than the front axle.

Dayton has perfected an automatic clearance adjuster for a hydraulic disc brake which is under development.

With reference to Mr. Hykes' paper, we have found that we can get organic linings which will peak at 800 F rather than the 400–500 F mentioned by the author.

We have an accelerated dynamometer test consisting of a series of panic stops. These tests have indicated that for a given lining drum life has a direct relation to weight. The curves show that an increase in drum weight from 75 to 78 lb increased drum life in the ratio of 8/13. Our tests have shown that much longer lived brake block materials are available on the market today than existed several years ago.

Heat checks are solely a function of temperature rise from the interface outward creating stresses beyond the elastic limit of the drum material.

THIS PAPER REPORTS on the present state of the art in the utilization of refractory metals for air frame and powerplant sheet metal components. By far the most promising of these metals to date is molybdenum.

The mechanical and physical properties of molybdenum are well-suited for high-temperature service. The combination of relatively high thermal conductivity, low thermal expansion coefficient, good specific heat, and a reasonably high emissivity of a coated surface make this material suitable for exterior surface application on severely aerodynamically heated components.

However, in its usable alloyed forms, molybdenum tends to behave in a brittle manner at room temperature, suffering from a high brittle-to-ductile transition temperature. Other unacceptable properties are the presence of laminations in the material, 45-deg preferred angle cracking, and difficulty of controlling interstitial alloying elements. The authors discuss each of these and the progress made in overcoming them.*

the Refractory

Alan V. Levy, Hughes Tool Co., Aircraft Division†
Saul E. Bramer, Hughes Tool Co., Aircraft Division†

AS FLIGHT VEHICLES have progressed from subsonic to supersonic aircraft and on to hypersonic and space vehicles, the materials of construction required to fabricate their components have also undergone a rapid progression to more and more complex forms. Today, the severe environments associated with modern flight have placed an ever increasing responsibility on the materials engineering portion of the total technology required to create weapons and space systems. High temperatures, severe chemical atmospheres, radiation, vacuums, and many other conditions not present in yesterday's aircraft must be carefully considered in selection of the materials of construction.

Probably the most important and difficult environmental component to be accounted for in any design is the elevated temperatures that occur as a result of high-speed flight. As operating temperatures of structural components approach or exceed 2000 F, a new materials technology is required to counteract successfully the weakening effects such high temperatures have on the commonly used materials. Metal alloys based on iron, nickel, and cobalt lose their usable strength at about 2000 F and actually begin to melt at temperatures slightly above 2000 F. The materials and design engineer must revert to materials systems heretofore not used in aircraft construction. The use of ceramics, ablation materials, complex cooling systems, and the refractory metals molybdenum, columbium, tantalum, tungsten, and possibly others are the means by which the designer accounts for the environmental temperatures that are present.

The extension of the metallic state into the temperature region above 2000 F is the most efficient, lightest weight, simplest construction method of creating structural components for this type of service. Fortunately, the refractory metals, primarily molybdenum, are in an advanced enough state of development to permit their utilization in the hottest regions of short sustained life vehicles. The use of ceramics, ablation plastics, and regenerative cooling systems have been successfully used in transient life applications. However, it is the refractory metals that primarily allow sustained life vehicles to be considered where + 2000 F service temperatures are encountered.

Material Comparisons

Molybdenum and its solid solution alloys form the vanguard of the refractory metals that will be used in advance in-the-atmosphere and space vehicles. Molybdenum is a unique material from a design application standpoint in that there is little need to be concerned about its ability to perform at elevated temperature service conditions, but there is great need for concern and extreme care in handling molybdenum components at no-load, ambient temperature conditions. Molybdenium, in its present state of development, is a metal that is not completely compatible with the ambient temperature environment. In its usable alloyed forms, it

† Formerly with Marquardt Corp.
* Paper presented at SAE National Aeronautic Meeting, New York, April 1, 1959.

development of

Sheet Metal Structures

tends to behave in a brittle manner at room temperature, suffering from a high brittle-to-ductile transition temperature. In fact, most of today's usable strength level refractory metal alloys behave in a brittle manner at room temperature, either in bending or impact. It is this single fact that will limit refractory metal applications for the present to special areas of a vehicle or powerplant where either by design or handling or both, the brittle characteristics of the material will not be detrimental.

The evidence of ductile behavior at temperatures below 0 F in molybdenum and especially columbium in the laboratory is the gleam of light for the future and one of the reasons for adapting present-day alloys to vehicle structure on a development basis. In the near future, completely ductile alloys will be available and the knowledge of lightweight, hot refractory metal structures gained by using reasonable care in handling today's room temperature brittle materials will be well worth the risks taken.

In particular, the rapid development of columbium-base alloys such as the 0.5% Zr, the 0.4% Zr, 40% Ta alloy and new General Electric alloys designated F44 and F48 is beginning to show specific compositions that are completely ductile at temperatures below room temperature. While the first two mentioned columbium alloys are considerably below the strength levels of the best molybdenum alloys and actually are only comparable to the strongest nickel-base alloy above 2000 F (Hastelloy W), the promise of higher strength, completely ductile materials in the future is great.

Of the four principal refractory metals that are available in sufficient quantity for structural applications, only molybdenum and columbium alloys are being considered for sheet metal components today. Tantalum, while the most ductile of the four and having the best overall fabricability, is not used because its strength-to-weight ratio is very low and its cost comparatively high. Tungsten is only now being investigated experimentally in sheet form and holds promise for use over 3000–3500 F. Below that temperature range, its strength-to-weight ratio is very low compared to molybdenum. Also, the alloy development status of tungsten is considerably behind that of molybdenum and columbium.

Table 1 is a comparison of the critical properties of the four principal refractory metal elements. The most promising alloy of each is used in the comparison to, in effect, give a quick review of the status of the refractory metals sweepstakes. It can be seen that the 0.5% Ti-molybdenum alloy is the

Table 1 — Critical Properties of Principal Refractory Metal Elements

Property	0.5% Ti Molybdenum	0.4% Zr-40% Ta Columbium	Commercially Pure Tantalum	Commercially Pure Tungsten
Melting Temperature, F	4,750	4,500[a]	5,425	6,150
Density (lb/cu in.)	0.368	0.368	0.600	0.697
Ultimate Tensile Strength in psi at Room Temperature	115,000	109,000	60,000	80,000
2000 F	74,000	29,000	19,000	34,000
2400 F	40,000	11,000	16,000	25,000
3000 F	10,300	—	12,500	20,000
Elongation at Room Temperature, %	10	45[a]	40	0
Bend Transition Temperature, F	100	Below −40	Below −40	—
Modulus of Elasticity	46×10^6	20×10^{6a}	27×10^6	59×10^6
Formability	Fair	Good	Good	Poor
Thermal Expansion at 2000 F in./in./F	3.4×10^{-6}	3.8×10^{-6}	3.6×10^{-6}	2.8×10^{-6}
Thermal Conductivity at 20 C cal/cm²/cm/C/sec	0.35	0.13[a]	0.13	0.40
Weldability	Fair	Good	Good	Poor
Coatings	Good	Good	Poor	Poor

[a] Extrapolated values.

Table 2 — Elevated Temperature Mechanical Properties of 0.5% Ti-Molybdenum

Tensile Properties

Test Temperature, F	Material Thickness, in.	Soak Time, min	0.2% Yield Strength	Ultimate Tensile Strength	Elongation, %
2400	0.060	0.5	33.5	39.9	5
2600	0.060	1.5	11.9	18.9	12
2600	0.040	1.0	11.4	16.2	10
2800	0.060	1.5	8.4	15.0	22
2800	0.040	2.0	8.7	12.4	10
3000	0.060	0.5	5.7	9.8	30
3000	0.040	1.0	6.7	10.8	20

Strain Rate: 0.001 in./in./sec to yield
0.01 in./in./sec to fracture

Creep and Stress Rupture Properties

Test Temperature, F	Stress, psi	Time (sec) to Obtain Creep Strain of 1.0%	2.0%	3.0%	4.0%	5.0%	Time to Rupture, sec	Elongation, %
2400	28,000	180	280	296	—	—	300	7
2600	10,000	37	82	122	266	—	272	8
2600	7,000	860	1800	2900	3810	4900	5400[a]	5.8%[a]
2800	7,500	57	180	336	505	663	1405	14%
2800	6,000	1474	2520	—	—	—	3300[a]	2.5%[a]
3000	4,500	56	113	170	214	246	290	12
3000	4,000	252	472	700	835	1060	1380	13
3000	3,500	238	720	—[b]	—[b]	—[b]	2883	16

[a] Test stopped at this time and elongation.
[b] Extensiometer slipped off specimen.

overall most promising material developed to date. Its primary shortcoming is its tendency to be brittle at room temperature, a very serious problem that could offset its other attractive properties in many kinds of applications.

Molybdenum Properties

Molybdenum has, by far, had the most extensive application in components and it is a detailed review of the experience that has been gained using this material in sheet form that will be presented in this paper. Fig. 1 shows the ultimate strength of as-rolled and stress-relieved commercially pure, 0.5% Ti alloy and 0.5% Ti, 0.07% Zr experimental alloy. The great promise of molybdenum as a high-strength structural material at temperatures above 2000 F is readily seen. The per cent increase in tensile strength for the first three alloys to come out of the laboratory is extremely promising.

Along with increased strength goes higher recrystallization temperatures. Commercially pure sheet can be recrystallized after 1–2 hr at temperatures as low as 1800 F. The 0.5% Ti alloy will not recrystallize until temperatures exceeding 2200 F have been reached. The 0.5% Ti, 0.07% Zr alloy will only crystallize above 2500–2600 F. Alloys still in the laboratory retain their cold-worked grain structure and corresponding higher strength to temperatures in excess of 3000 F. Since no dispersion hardening mechanisms have yet been developed for molybdenum alloys, the retention of cold-worked strength is an important consideration.

As noted in Table 1, the high modulus of elasticity of molybdenum is an attractive property, especially where thin sheet metal structure is desired. An additional attractive property of molybdenum alloys is their excellent creep strength. Table 2 shows some of the short-time tensile and creep properties of 0.5% Ti-molybdenum alloy at temperatures from 2400 to 3000 F. It can be seen that for the low-load, high temperature environment for limited time periods of atmosphere re-entry or high-altitude hypersonic flight, molybdenum is well-suited. For longer life applications, present molybdenum alloys are well-suited for service at temperatures in the 2000–2500 F range. Fig. 2 shows the creep and stress rupture properties of the 0.5% Ti-molybdenum alloy at 2400 F.

The physical properties of molybdenum are well-suited for high-temperature service. The combination of relatively high thermal conductivity, low thermal expansion coefficient, good specific heat, and a reasonably high emissivity of a coated surface make this material well-suited for exterior surface application on severely aerodynamically heated components. Table 3 lists some of the more important physical properties.

Table 3 — Physical Properties of Molybdenum

Atomic Number	42
Atomic Weight	95.95
Lattice Type	Body centered cubic
Density (lb/cu in.)	0.368
Melting Point	4720 ± 18 F
Boiling Point	8670 F
Specific Heat (cal/gm/C)	
50 C	0.06
500 C	0.07
Heat of Fusion (Btu/lb)	126
Heat of Sublimation (Btu/lb)	2880×10^3
Heat of Combustion (Btu/lb)	3261.6
Coefficient of Thermal Conductivity (Btu/ft²/ft/hr/F)	
Room Temperature	76
400	71
800	66
1200	63
1600	60
1800	58

Mean Coefficient of Thermal Expansion (in./in./F × 10^{-6}) Temperature Range	Mean Coefficient
32/200	2.67
32/1000	2.98
32/1600	3.21
32/2000	3.36
32/2600	3.59
32/3200	3.81

Probably the best way to present the darker side of the molybdenum picture is to discuss the present quality of incoming heats of alloy sheet material. Marquardt Aircraft Co. has had experience with sheet material from over 60 different heats of 0.5% Ti alloy ranging in gage from 0.010 to 0.080 in. and in sizes up to 36 × 60 in. From a detailed quality control analysis of this material has come a "feel" for the "state-of-the-art" that has both bright and dismal aspects, but an overall glimmer of better things to come.

First, the good side. The uniformity and level

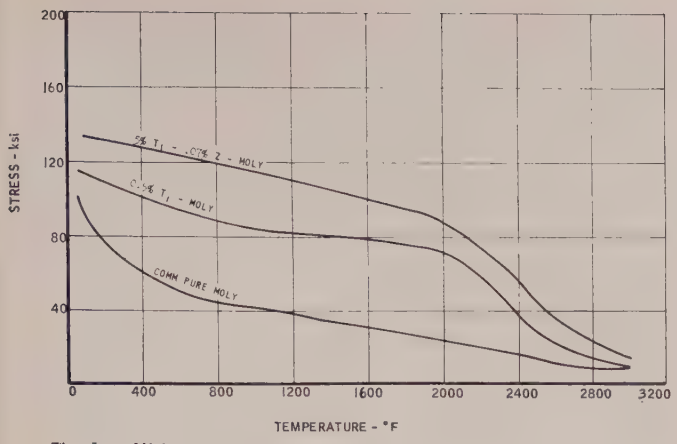


Fig. 1 — Ultimate tensile strength versus temperature for various refractory metals

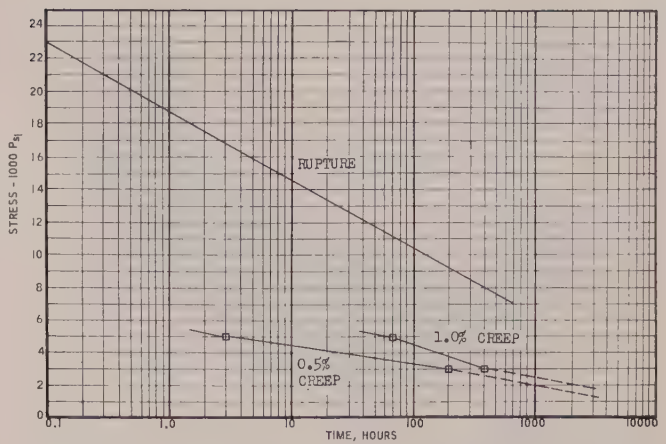


Fig. 2 — Creep and stress rupture curves for 0.5% titanium-molybdenum alloy at 2400 F in recrystallized condition

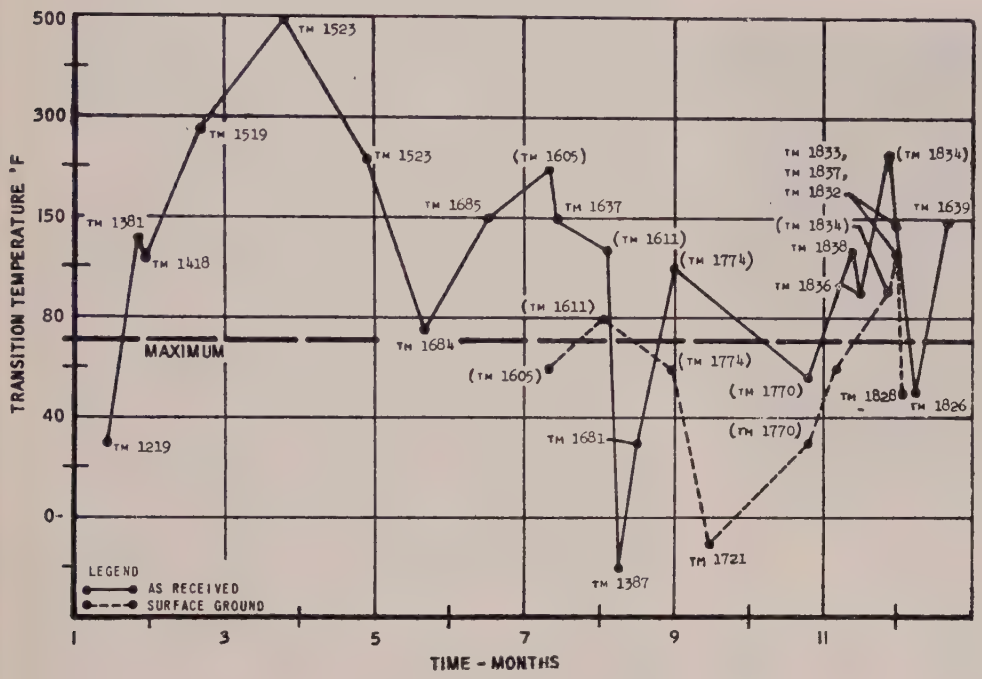


Fig. 3 — Bend transition temperatures for as-received and surface-ground 0.5% titanium-molybdenum alloy sheet for 1958

of the following properties are considered acceptable and on a par with the general metal producing industry:

1. Ambient tensile strength and elongation.
2. Chemistry (except for somewhat questionable level of interstitials such as oxygen and nitrogen).
3. Surface condition and flatness (after surface grinding 0.002–0.005 in).
4. Formability and weldability.

The acceptable status of these properties has been reached only in the past year. Prior to that time, almost any level of a given property could be expected from heat to heat and even from sheet to sheet within a heat. This improvement alone is gratifying.

The unacceptable level properties, the properties that one must live with for the present if he wants to develop lightweight, efficient structures for 2000 F and above operation, are listed below. The causes of some of these problem areas are subtle metallurgical factors possibly going back all the way to the ore or reduced powder. Others are primarily due to conversion into shapes, extrusion, rolling. There are no answers or solutions to eliminating the undesirable nature of these properties to date and it is only with a concerted effort on the part of industry and the government together that these problems will be solved.

Brittle-Ductile Transition Temperature

The most critical shortcoming of today's molybdenum alloys is the brittle behavior they exhibit at the lower temperatures. While certain laboratory and a few production heats have exhibited tensile and bend transition temperatures below 0 F, the vast majority of production sheet metal has had a bend transition temperature at or above room temperature. The impact transition temperature of molybdenum alloys has consistently been in the 300–500 F range. The ability of molybdenum sheet

Fig. 4 — 0.5% titanium-molybdenum spun cup

Fig. 5 — Impact test of molybdenum weld panel U14

to absorb low energy level impacts at room temperature (such as occur in shop handling) has varied from heat to heat and does not directly reflect the transition temperature in either bending or impact. A determination of "how brittle is brittle?" is presently underway.

Fig. 3 shows the variation in bend transition temperature from heat to heat over a relatively large number of heats. The curve speaks for itself. One of the interesting aspects of the bend transition property, as shown by the dotted lines on the curve, is the significant improvement that can be made by surface grinding off contaminated material. However, it is only when the as-rolled surface is quite contaminated, resulting in a bend transition temperature above about 200 F, that removing this surface is beneficial.

The problem of improving the brittle-ductile transition temperature property of molybdenum alloys is a complex one that must be approached by an across-the-board development program aimed at every phase of production from reducing the ore to rolling the sheet bar into sheet.

Laminations

The other critical problem confronting the user of molybdenum alloy sheet metal is the tendency to form and the actual presence of laminations in the material. Available rolling mill equipment for producing steel sheet product is limited to metal working temperatures of about 2300 F. Since most materials are best reduced by hot rolling at 70–80% of their melting temperatures, equipment for rolling at metal temperatures to 2300 F is completely adequate for all materials in use today except the refractory metals. Since molybdenum alloys do not recrystallize until temperatures of 2300–2400 F are reached, all reduction of molybdenum on today's rolling mills is done cold or below the recrystallization temperature. This factor has a major effect on the material being rolled. Because it is being rolled at less than 50% of its melting temperature, large differential shear stresses exist between the layers of sheet passing through the rolls, inducing shear separation and areas of high residual shear stress primarily along the flat grain boundaries that exist in the plane of the sheet. Also, the inability of laps and seams present from the initial operations to weld together as rolling progresses contributes to laminations. Any load applied to the finished sheet that has a force component that adds to these highly stressed areas or acts upon actual separations can open up laminations in the material. Fig. 4 is a manually spun cup of 0.5% Ti-molybdenum alloy that shows these laminations to a marked degree. In general, the degree of laminations present in a sheet is apparent only if the material is sheared or broken and show up as a micaceous layering along the fractured surface. However, in some instances, the lamination has been so extensive that the sheet actually consisted of two separated sheets over large surface areas.

The problem of laminations is primarily tied to the conversion process into sheet form. Rolling at true hot-rolling temperatures should markedly improve this problem area. Making lower per cent reductions between anneals or stress reliefs on hot cold-worked material should help the presently produced molybdenum alloys. It has not been determined whether minor laminations in a sheet are detrimental to its service performance. Evidence on both sides exists. Small-size laminations do not effect fabricability other than fusion welding where they can not be tolerated. Laminations certainly are not desirable. However, it has been found that they can be lived with, temporarily, provided they are not excessive.

45-deg Preferred Angle Cracking

Molybdenum sheet metal exhibits an anisotropic property in its tendency to propagate cracks at very fast rates in a direction that is 45 deg from its rolling direction. This behavior is extremely intermittent in nature, but when it occurs, it is particularly harmful because of the apparent extreme ease with which the cracks can traverse great distances. The differences between transverse and longitudinal and even 45-deg tensile properties give no clue as to the presence or cause of the 45-deg cracking. It has been observed that this type of cracking requires a high stress concentration such as exists at a notch or a short transverse crack in a fusion weld to be initiated. Fig. 5 shows an extreme case of 45-deg cracking initiated from a transverse crack in a fusion weld.

Since it is not known exactly what causes the 45-deg cracking tendency, its elimination will have to await further research work. However, in practice, it can be prevented from occurring by not introducing sharp stress concentrations into parts.

Recrystallization

The properties of the solid-solution-type alloys presently being produced in refractory metals are enhanced only by introducing cold work into the sheet during its last rolling reduction and after the final annealing treatment. The ultimate tensile strength of the 0.5% Ti-molybdenum alloy is increased from 33,500 to 74,000 psi at 2000 F by cold working. If the recrystallization temperature is exceeded either in service or in processing such as in a coating operation, this marked strength increase is lost. Because of the varying amounts of cold reduction used in the final pass in rolling today's sheet, recrystallization temperatures of the 0.5% Ti alloy can vary over a range of several hundred degrees in the temperature range from 2200 to 2450 F. In a few cases, the annealing temperature has actually dropped into the range of the coating processing and resulted in partial recrystallization prior to service.

It is important to the user that the producer standardize his rolling practice to stabilize the recrystallization temperature of existing alloys at as high a level as is possible. The creation of new alloys with recrystallization temperature above 3000 F is an important future step that is being taken in the laboratory.

Control of Interstitial Alloying Elements

The problems involved in controlling the interstitial alloying elements oxygen and nitrogen in molybdenum alloys is complicated by two facts. It is very difficult to determine accurately the level of oxygen and nitrogen present; and, at the lower quantities, it is difficult to set an acceptable minimum that is based on a knowledge of the detrimental effects that occur. It is relatively easy to determine that 0.1% oxygen severely embrittles molybdenum. However, what effect 0.005% O_2 has on the material versus what effect 0.002% O_2 has is beyond the state of the art. Therefore, little can be said concerning this problem other than the lowest level interstitials possible is desirable.

Fabricability

The fabrication of powerplant and missile components from commercially pure molybdenum and 0.5% Ti-molybdenum alloy has been accomplished by all of the commonly used fabricaiton techniques. Because of the lack of room-temperature ductility, all handling, forming, and joining must be done at elevated temperatures. Shop personnel must be trained to handle the material properly and to respect its ambient temperature limitations. The result of the extensive fabrication and coating development work culminated in the fabrication of an all molybdenum riveted and oxidation protection coated tailpipe (Fig. 6).

Sectioning

Molybdenum sheets can be sectioned by shearing and sawing when the proper precautions are taken. Machine shearing is done by heating the entire sheet to several hundred degrees Fahrenheit and then

Fig. 6 — 0.5% Ti-molybdenum riveted combustion chamber and exit nozzle

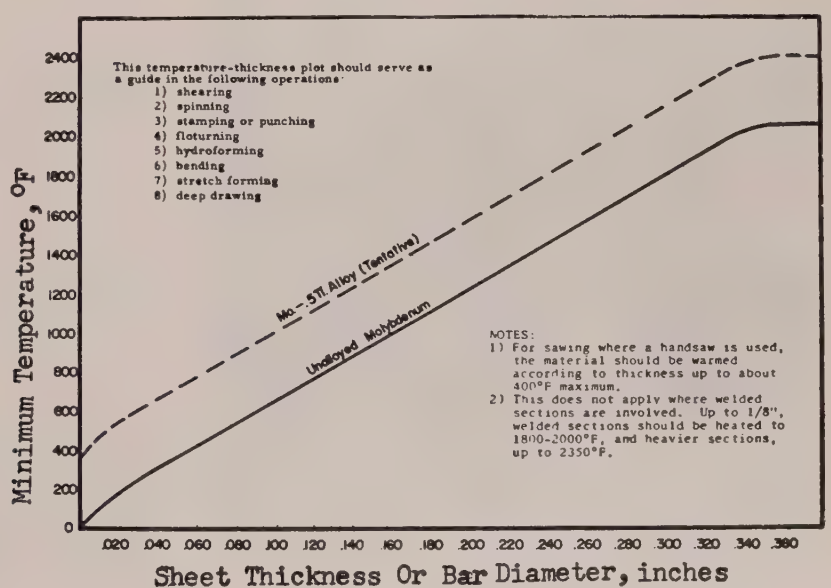

This temperature-thickness plot should serve as a guide in the following operations:
1) shearing
2) spinning
3) stamping or punching
4) floturning
5) hydroforming
6) bending
7) stretch forming
8) deep drawing
Mo - 5Ti Alloy (Tentative)
Unalloyed Molybdenum
NOTES:
1) For sawing where a handsaw is used, the material should be warmed according to thickness up to about 400°F maximum.
2) This does not apply where welded sections are involved. Up to 1/8", welded sections should be heated to 1800-2000°F, and heavier sections, up to 2350°F.
Minimum Temperature, °F
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
.020 .040 .060 .080 .100 .120 .140 .160 .180 .200 .220 .240 .260 .280 .300 .320 .340 .360
Sheet Thickness Or Bar Diameter, inches

Fig. 7 — Suggested minimum working temperatures for molybdenum and Mo-0.5% Ti shapes

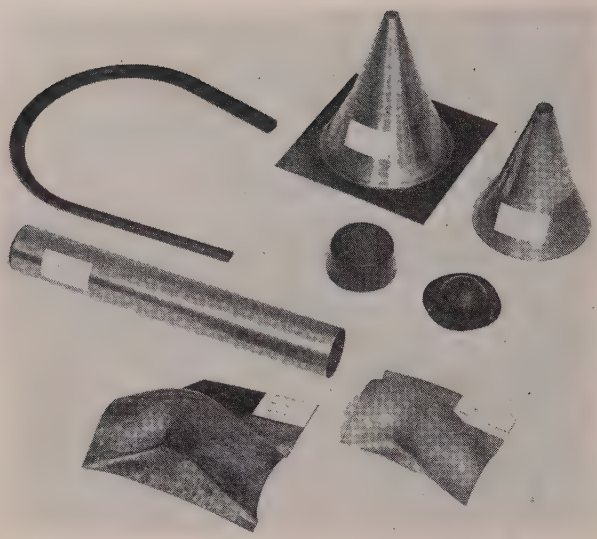


Fig. 8 — Molybdenum formed shapes

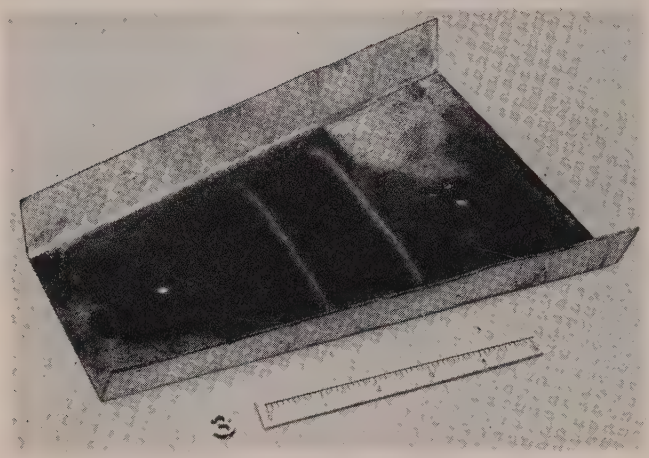


Fig. 9 — 0.5% Ti-molybdenum alloy drop hammer formed part

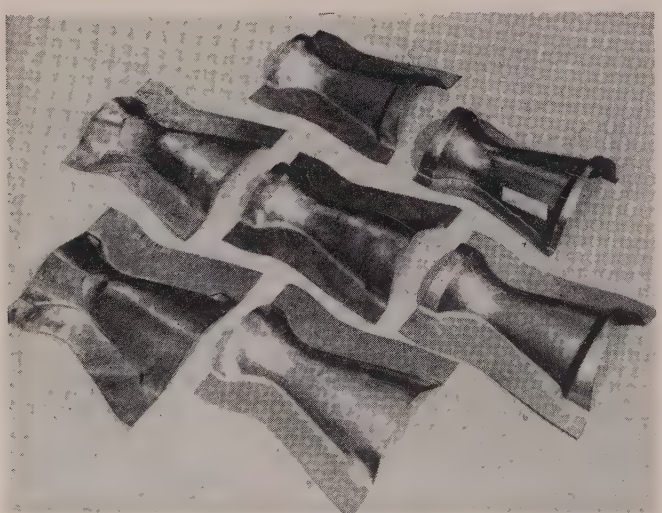


Fig. 10 — All stages — combined picture of molybdenum tailpipe half-section

Fig. 11 — Spinning of 1% Ti-molybdenum for tailpipe components

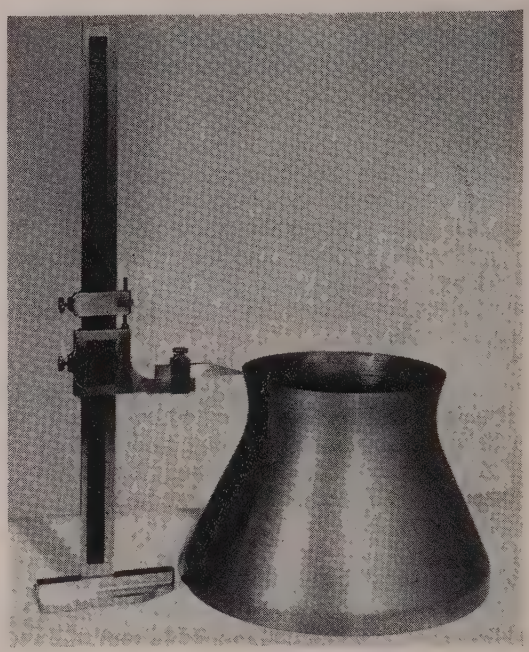


Fig. 12 — Single piece, manually spun nozzle section

locally torch heating the area to be sheared to between 400 and 1000 F. A tempil-stick is adequate for a temperature measurement. In hand shearing, such as on a Beverly Hand Shear, the material should be locally heated to a dull red, which is about 1000 F, immediately ahead of the shear blade before cutting. By heating to these temperatures, the sheet will be well above its brittle-to-ductile transition range, cracking, and opening up of laminations will be eliminated. By using a fine tooth blade at very high speeds on a band saw, flat patterns and formed molybdenum sheet sections can be cut. The creation of a heat zone immediately ahead of the saw blade as a result of friction caused by the high speed eliminates cracking in the material. Care should be taken to join the saw blade properly to avoid any shock associated with misaligned joints. An abrasive cutoff wheel with coolant also can be used to section molybdenum sheet.

Forming

The basic technique required to form shapes is primarily dependent upon the ability to control forming temperature as related to forming speeds and sheet thickness. Fig. 7 gives the suggested minimum working temperatures for molybdenum and 0.5% Ti-molybdenum alloy. Although the above relationship can be used as a guide, in some cases

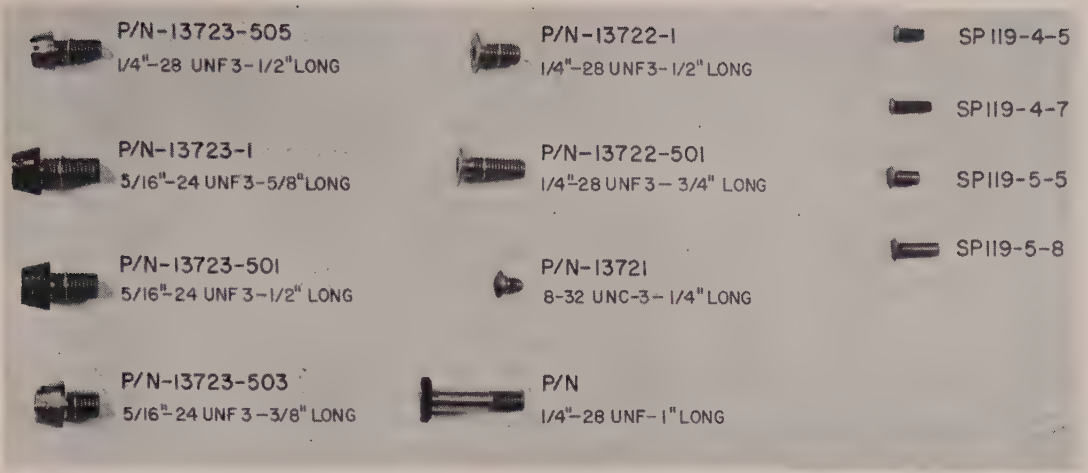


Fig. 13 — 0.5% Ti-molybdenum alloy fasteners

experience and "feel" for the material must be utilized to form a part successfully. Because of the high cost of the material, the chipless operations such as extruding, die forging, radial draw forming, manual spinning, hydroforming, impact forming and spin forging (flo-turn or hydrospin is synonymous with spin forge) should be utilized as much as possible.

Fig. 8 illustrates different shapes of commercially pure molybdenum and 0.5% Ti-molybdenum alloys formed by various techniques. At the top of the picture can be seen spin-forged nose cones and a stretch-wrapped section used for a structural ring. Beneath the nose cones are a cup which was manually spun on a spinning lathe and a hydroformed hemisphere. The seamless tube was produced by spin forging a heavy wall extruded preform blank.

At the bottom of the picture are two impact elbow half-sections. The nose cones, seamless tube, spun, cup, and hemispheres were formed with the material at approximately 600 F. The radial draw formed section was made with the material at 250 F and the elbows were impact formed at 1200 F. Air frame channel sections with built-in joggles and stiffeners have been impact formed at 1200 F (Fig. 9).

Molybdenum alloy sheet can be readily formed at elevated temperatures into complex configurations. Figs. 10–12 show two different approaches of producing a contoured nozzle section. Fig. 10 shows impact formed nozzle half-sections in various forming stages with the finished part in the upper right-hand corner. To eliminate longitudinal joints, manual spinning techniques were developed for producing seamless nozzle segments. Fig. 11 shows the preformed cups and two finished sections made from it. Fig. 12 shows a single piece, manually spun nozzle section.

In general, it has been found that as long as molybdenum sheet is formed at elevated temperatures, it is as formable as any other sheet metal alloy and can be formed into complex shapes accurately.

Joining

Sheet metal assemblies incorporate joints of several types in their fabrication. Molybdenum alloy mechanical fasteners (screws, bolts, and rivets) have been produced and welding techniques for resistance, fusion, and flash welding are being developed.

Mechanical Fasterners — An accelerated mechanical fasteners development program was initiated early in the molybdenum application program to fabricate rivets, bolts, and screws and to use them in joining molybdenum components. Fig. 13 shows the 0.5% Ti-molybdenum alloy fasteners that have been developed. All fasteners are hot headed. Fig. 14 is a photomacrograph of a rivet and Fig. 15 is a photomicrograph of a bolt showing the flow lines which result from upsetting the heads.

To install rivets successfully, they must be heated to 1000–1200 F. Care should be taken to keep the

Fig. 14 — Hot-headed 0.5% Ti-molybdenum rivet

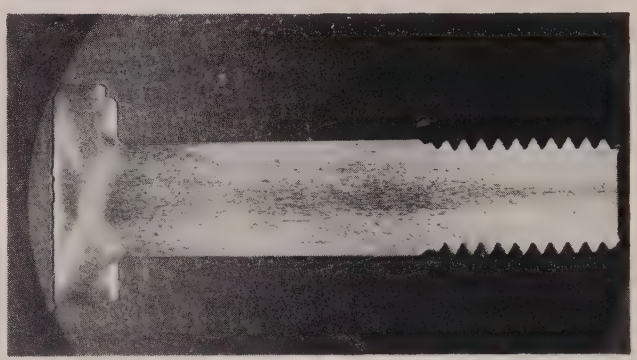


Fig. 15 — Hot-headed 0.5% Ti-molybdenum bolt

metal component being joined at several hundred degrees Fahrenheit to insure that the base metal will not crack when the rivet is squeezed. Bolts can be installed at room temperature; however, care must be taken that bolt holes are aligned perfectly because any excess torque due to interference will shear the bolt before the threads will deform (Fig. 16). Bolts and screws can be coated for oxidation resistance prior to assembly, making allowance for the coating thickness. Riveted joints have been successfully oxidation protection coated after assembly.

Resistance Welding — The development of spot-welding techniques has proceeded very satisfactorily. The principal variables investigated were:

1. Evaluation of methods of surface preparation.
2. Determination of the proper type of spot-welding electrode.
3. Determination of welding techniques for producing optimum spot welds.

Electrolytic polishing, chemical cleaning, and mechanical cleaning, plus a solvent wipe were the cleaning methods investigated. The chemical method of cleaning appears to be the most satisfactory method because of the relative ease of the cleaning operation. The elaborate setup required for electrolytic polishing makes this method impractical. Mechanical cleaning plus solvent wipes proved unsatisfactory because the oxide film could not be removed without removing considerable amounts of material.

The best electrodes to use are those which have a very high conductivity and are not the current collector type. Current collector types have face-heating tendencies and make the electrodes stick to the base metal causing copper diffusion into the base metal.

To produce optimum spot welds, molybdenum's characteristics of having a high thermal conductivity and high brittle-to ductile transition temperature must be accounted for. The high thermal conductivity dissipates the heat rapidly away from the joining area and the use of long weld current times results in nonbonding of the faying surfaces. Short cycling, that is, the welding amperage must be brought in rapidly and discontinued almost immediately, must be used to produce spot welds. To avoid surface cracking, the preheat control must be used to raise the temperature of the material above the transition temperature prior to applying the welding current. Immediately after welding, the post-heat circuit and forging actions are initiated. The post-heat keeps the material above the brittle-to-ductile transition temperature and forging aids in grain refinement. Fig. 17 is a photomicrograph of a weld nugget, and Fig. 18 shows a 0.5% Ti-molybdenum spot weld bent over a 1T radius at 500 F. Resistance welds are completely ductile above the joints transition temperature. As experience has been gained the degree of brittleness in spot-weld heat-affected zones has markedly decreased. Nuggets exhibit some metal flow in shear tests conducted at room temperature. Shear strengths of 900–1000 lb at room temperature are obtained when joining 0.040–0.040 in. thick material and 1800–1900 lb of shear strength are obtained when joining 0.056–0.056 in. thick material.

Fusion Welding — Molybdenum has been successfully fusion welded in the open air using local protection without the need for inert atmosphere boxes. Inert atmosphere (either helium or argon or a combination of both gases) gas blankets on the top and bottom of the welds are supplied by modifying the torch and backup bars to receive an inert gas supply. Enough exploratory work was done on crude equipment to formulate what refined equipment requirements were necessary to fusion weld molybdenum successfully so that performance specifications for longitudinal and circumferential positioners could be created. Fig. 19 shows the longitudinal positioner and Fig. 20 shows the circumferential positioner that were developed specifically for welding molybdenum alloys. Some of the unique features of the longitudinal welder are:

1. The welding head travel is accomplished by a precision lead screw with the head guided on hardened ground rails to insure consistent and constant welding head speeds. This eliminates the possibility of any intermittent slowing down of the welding head as occurs when a rack and pinion arrangement is used for head travel. The extreme fluidity

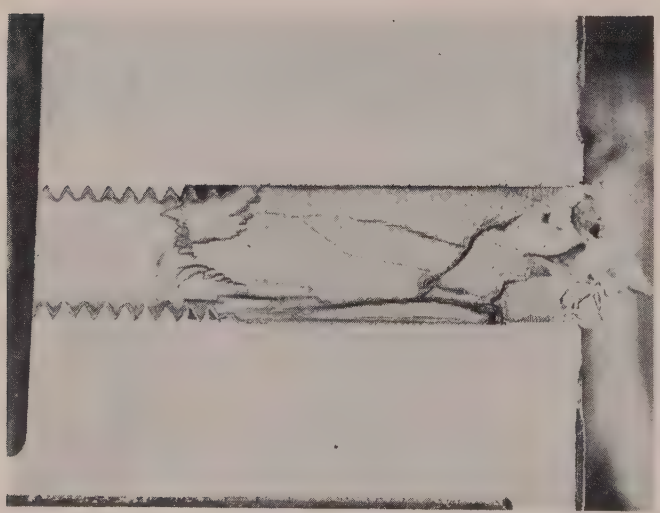


Fig. 16 — Deformation of imperfectly aligned bolt

Fig. 17 — Photomacrograph of molybdenum spot weld

and high melting temperature of molybdenum requires that constant welding speeds be used to prevent blowing holes in the material.

2. The back-up mandrel is so designed that the maximum deflection is less than 0.005 in. when the full pressure on the hold-down bars is applied. This feature is necessary to maintain a constant arc length to insure constant arc voltage over the full length of the weld at the very high welding speeds up to 100 in./min.

3. The hold-down bars can be accurately positioned to insure that the same weld heat will be conducted away from each side of the weld joint during welding so that the weld bead width remains constant.

4. The hold-down pressures are very uniform and accurate to prevent weld cracking.

5. Heaters are installed in the back-up bar to maintain preheat temperatures to better than ± 5 F in the part.

The circumferential positioner has a variable rotational speed control which is precisely constant at any speed setting. The back-up mandrels for the circumferential positioner must be able to preheat the part to better than ± 5 F and maintain intimate contact with the part over the entire circumference.

A program has been in progress on this new automatic equipment to develop sound, reproducible molybdenum weldments in heavy sheet with the lowest possible bend transition temperature. The effect of the following variables on the weld properties have been considered:

1. Tooling requirements.
2. Material analysis.
3. Material thickness.
4. Surface preparation.
5. Joint fit up.
6. Inert gas analysis
7. Preheat temperatures.
8. Effect of weld grain size and grain size gradient.
9. Effect of weld grain orientation.
10. Optimizing weld schedules.
11. Effect of removing weld surfaces.
12. Effect of stress relief treatments.
13. Effect of cold-work treatments.
14. Anisotropic properties of the material.

Molybdenum weldments are very sensitive to any restraints imposed on the joint. The correct combination of the hold-down bar pressure and the part preheat temperature can eliminate cross-cracking in molybdenum weldments. The following summary gives the results of the preheat investigation:

1. No preheat—38 cracks in 3 welds.
2. 150–200 F preheat—5 cracks in 8 welds.
3. 300 F preheat—no cracks in 9 welds.

The 300 F preheat also resulted in decreasing the

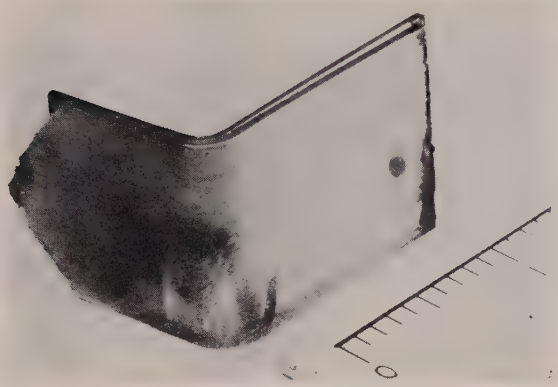


Fig. 18 — 0.5% Ti-molybdenum spot-welded sheet, bent over 1T radius at 500 F

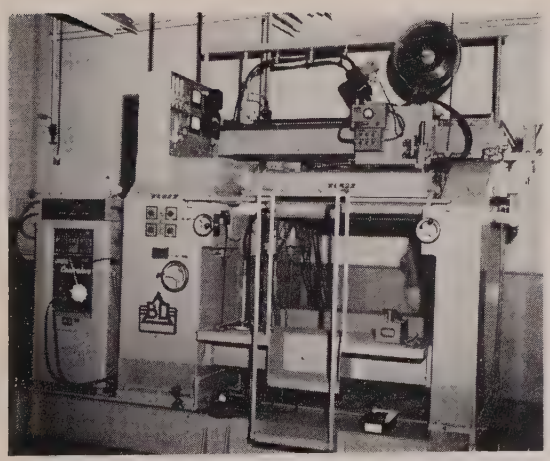


Fig. 19 — Automatic fusion welding machine (WL622) longitudinal type

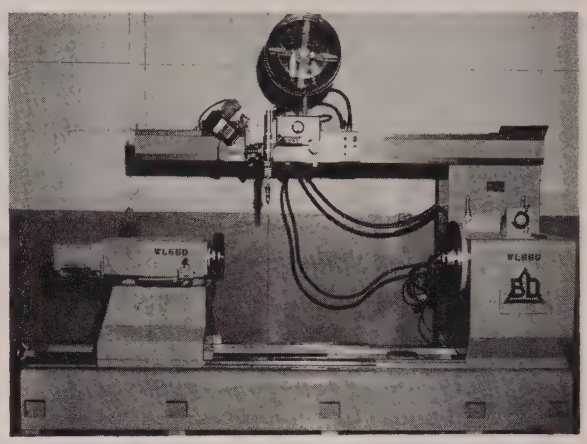


Fig. 20 — Automatic fusion welding machine (WL660), circumferential type

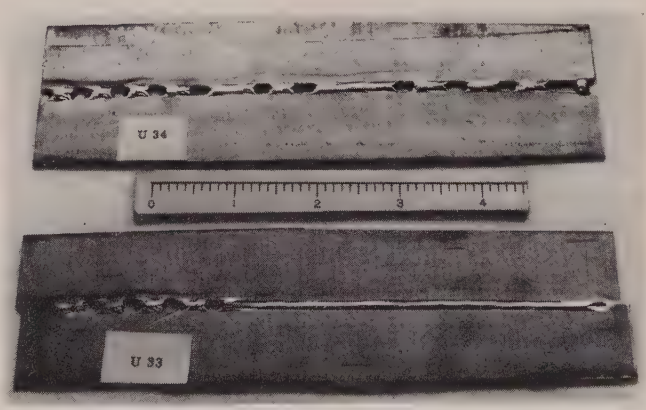


Fig. 21 — Sensitivity of molybdenum welds to preheat temperature differential of 25 F

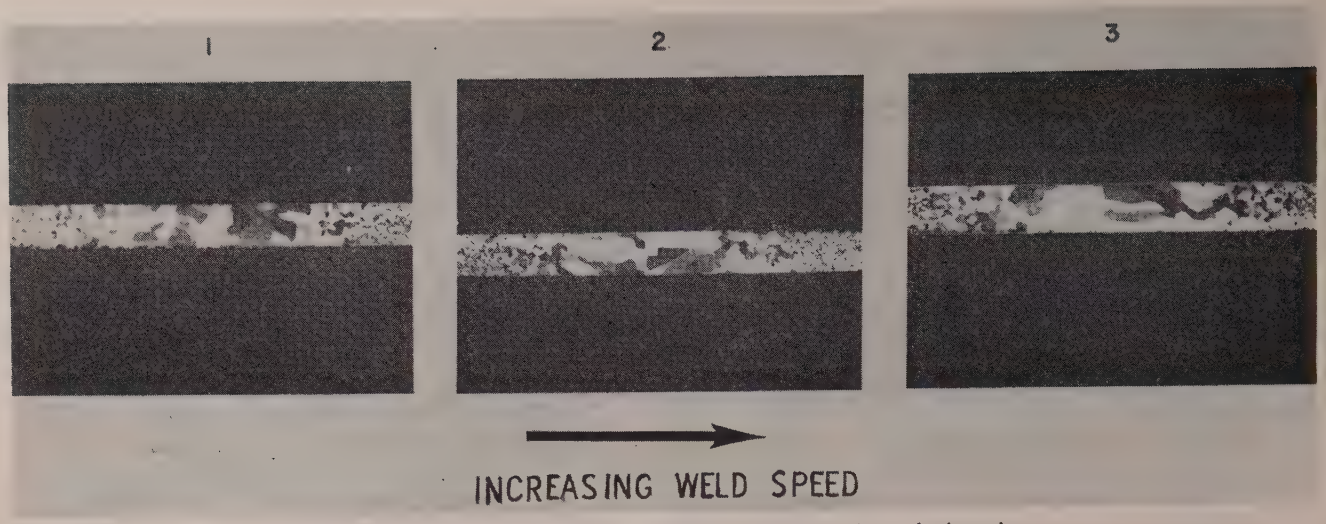


Fig. 22 — Effect of welding speed on molybdenum weld grain size and orientation

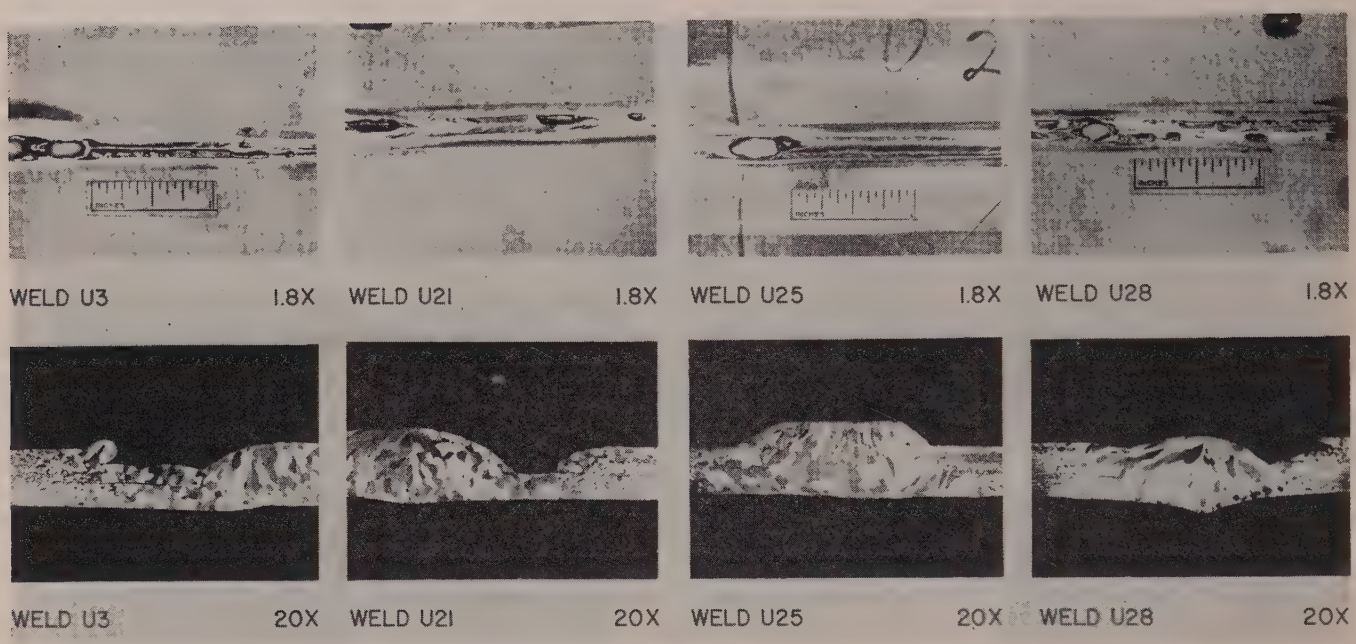


Fig. 23 — Evidence of laminations in molybdenum welds

welding current by 80 amp under that welding current used to weld specimens with no preheat. When all the welding conditions are constant, a variation of only 25 F in preheat temperature results in holes being blown the full length of the weld bead (Fig. 21) because of the sensitivity of molten molybdenum to temperature. Therefore, it is desirable that preheat temperature variations of ± 5 F be maintained to produce sound reliable weldments.

Another variable that affects the weld bend transition temperature is the welding speed. An increase in welding speed can decrease the bend transition temperature by 200 F. Fig. 22 shows the effect of welding speed on the grain size, grain size gradient, and grain orientation. The grain size gradient is decreased with increasing weld speed as well as the grains in the center of the weld bead becoming more favorably oriented to reduce the bend transition temperature.

It has also been determined that a post-weld stress-relieving treatment can reduce the weld bend transition temperature. A stress relieve at 1700 F for 4 hr reduces the bend transition temperature of weldments by 25–50 F. However, a 1500 F stress-relieving treatment for 4 hr has no effect on the weld bend transition temperature.

A study was made to correlate the bend transition temperature of recrystallized material with that of weldments containing recrystallized grains in the weld zone. To date, the best welds produced have a bend transition temperature of 175 F with no ambient temperature ductility improvement in sight. It was thought that this high bend transition temperature may be the fault of recrystallization in the weld zone rather than contamination. Table 4 gives the results of the recrystallization property study.

It can be concluded that with the base material in its present state of development, molybdenum weldments would never have an ambient bend transition temperature because recrystallized material whether it be in base metal or in a weld with a similar grain size has approximately the same level of increased bend transition temperature. This is apparently due to the recrystallized condition of the material rather than to any other factor. The increased bend transition temperature of the weld over that of the fully recrystallized base metal is due to oxygen contamination of the weld.

Two of the unfavorable characteristics mentioned earlier of today's as-received material, namely laminations (planes of weakness) and 45 deg cracking, are causes for greater concern in weldments than the bend transition temperature problem. Fig. 23 shows how laminations in the base metal can result in metal expulsion and the parting of these planes of weakness when heat is applied to the joint line during welding. Directional embrittlement, 45-deg angle cracks, originating from stress concentrations as a result of imperfections in a weld bead can be vividly seen in Fig. 5. As the state of the art progresses to the point where high-quality material can be produced without these defects, consistently sound weldments will be possible. Using today's material, careful selection of the individual sheets is necessary where fusion welding is required if sound, reliable welds are to be produced.

Flash Welding — Satisfactory results have been obtained in the flash welding of molybdenum bar stock. Fig. 24 shows flash welds in 0.5% Ti-molybdenum bar stock having a cross-section of $\frac{1}{4} \times 1$ in. Fig. 25 is a photomacrograph of the molybdenum flash weld showing the flow lines as a result of the upset and a weld joint without a recrystallized weld zone. At room temperature these welds had a 0.2% offset yield strength of 89,200 psi, an ultimate tensile strength of 99,800 psi, or an 85% ultimate joint efficiency, and an elongation of 6%.

Table 4 — Comparison of Commercially Pure Molybdenum[a] Recrystallized Base Metal Bend Transition Temperature with Weldments Having Approximately Same ASTM Grain Size

Specimen and Heat-Treatment	Microstructure Condition and ASTM Grain Size	Bend Transition Temperature F
As-received	Cold-worked and stress relieved	−45
2000 F/1 hr — argon	Partially recrystallized — ASTM No. 7	Room Temperature
2100 F/2 hr — argon	Fully recrystallized — ASTM No. 6	Room Temperature
2200 F/4 hr — argon	Fully recrystallized — ASTM No. 5½	100
2250 F/6 hr — argon	Fully recrystalized — ASTM No. 5	125
Weld	Surface ground — ASTM No. 3	225
Weld	Surface ground and stress relieved 1700 F/4 hr — argon — ASTM No. 3	175

[a] All material from same sheet.

Fig. 24 — 0.5% titanium-molybdenum alloy flash welds before and after flash removal (heat TM 1215)

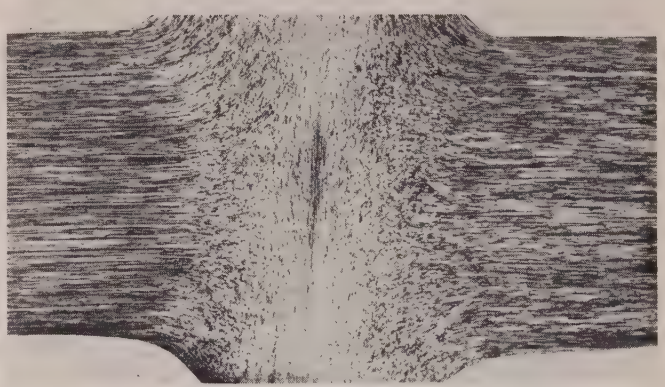


Fig. 25 — Photomacrograph of molybdenum flash weld showing weld and flow lines

Machining

The machining of molybdenum is not too difficult. For most machining operations, the material cuts like the stainless steels and a carbide-tipped tool should be used. When sheet edges are machined to prepare them for welding, mild-steel bars should be clamped on each side of the sheet to prevent chipping and separation of the planes of weakness. In the chasing of threads, the material machines like cast iron in that powdery, discontinuous chips are produced. Fig. 26 shows American Standard and Acme threads machined in molybdenum tubing. For drilling and boring operations, a chlorinated cutting oil should be used so that the chips can be continuously flushed from the hole because their abrasiveness can significantly decrease tool life.

Coatings for Oxidation Resistance

The greatest handicap of all the refractory metals (molybdenum, tungsten, tantalum, columbium, and the rest) is their inability to resist oxidation at elevated temperatures. The refractory metals, other than molybdenum, form an oxide scale on the surface when exposed to an oxidizing atmosphere at elevated temperatures. However, above about

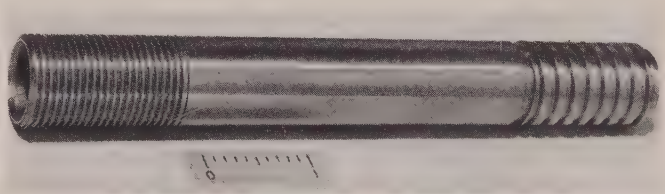


Fig. 26 — Machining test of 0.5% Ti-molybdenum tubing showing standard and Acme threads

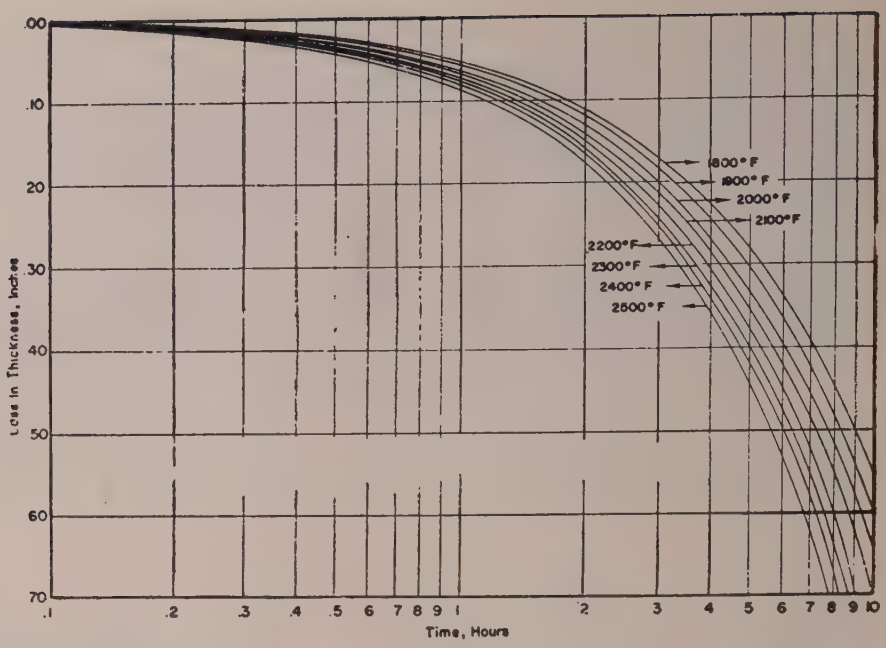


Fig. 27 — Total loss in thickness from oxidation of molybdenum sheet exposed to air on both sides

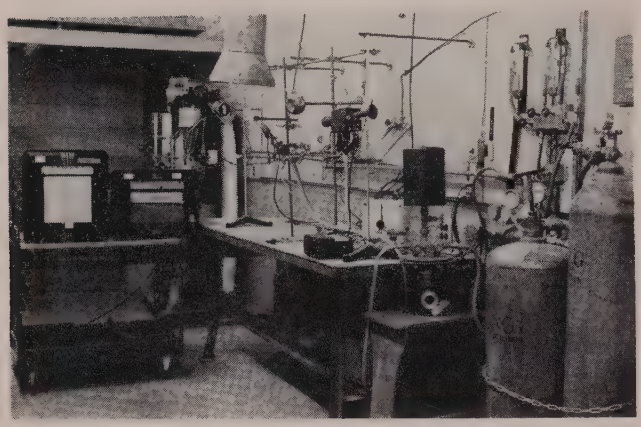


Fig. 28 — Thermal drop test setup

1000 F molybendum volatilizes as molybdic oxide (MoO_3) and the rate of this phenomenon increases rapidly as the temperature increases. At about 1400 F the rate of volatilization equals its rate of formation. Fig. 27 shows the loss in thickness from oxidation of molybdenum sheet exposed to air on both sides at elevated temperatures. At 1800 F, the oxidation rate for molybdenum in slowly flowing air is greater than 0.022 in. per hr per surface. To overcome this serious disadvantage to the use of molybdenum as a structural metal at elevated temperatures much research has been done on the development of coatings for the oxidation protection of molybdenum. Hot dip, paint and sinter, flame spraying, electroplating, electroless plating, electrophoretic deposition, vapor phase deposition, flame plating, cementation, and cladding coatings have been developed. This paper will cover some of the coatings evaluated at Marquardt Aircraft Co. Fig. 28 shows the thermal drop equipment for testing ceramics and oxidation protection coatings for refractory metals.

Chromalloy Corp. W-2 Coating (Cementation)

The Chromalloy W-2 coating is the best all-around coating tested to date. This coating is applied by packing the part in a powdered mixture in a retort, sealing the retort and heating at 2000 F for a number of hours. The powder generates a gas which diffuses into the molybdenum and results in a thin (approximately 0.002 in thick) alloy case with a surface finish as smooth as the original base-metal. The advantages of the W-2 coating include: unlimited throwing power, uniform, thin case depth on all exposed surfaces, excellent reproducibility, diffusion temperature below the recrystallization temperature of the base metal, resistance to thermal shock, and metallic-like surface finish.

A few of the tests which have been conducted on the W-2 coating are:

1. A coated specimen has been oxyacetylene flame heated to 2250 F, held at temperature for 5 min, and forced aircooled for a total of 20 cycles without failure.

2. Tensile and creep rupture tests have been conducted with coated specimens at temperatures to 3000 F, times to 90 min and elongations of about 7% at 3000 F before molybdic oxide vapors were observed with subsequent coating break-down.

3. A sample was bent (after coating) to a 15 deg angle and heated in a furnace at 2350 F for 400 hr without failure.

4. A coated panel was bent 10 deg and tested using a high velocity combustion element as the burner. The coated panel was heated to 2740 F for 3 min and then the temperature was raised to 3420 F for 72 sec without coating failure. Fig. 29 shows a W-2 coated panel before and after the exposure.

5. A metallographic examination of a molybdenum riveted joint specimen, which contained a fabrication crack and was assembled complete prior to coating revealed complete coverage of the joint and diffusion of the coating inside the cracks. The multirivet panel was subjected to a temperature of 2700 F for 20 min with only minor defects resulting

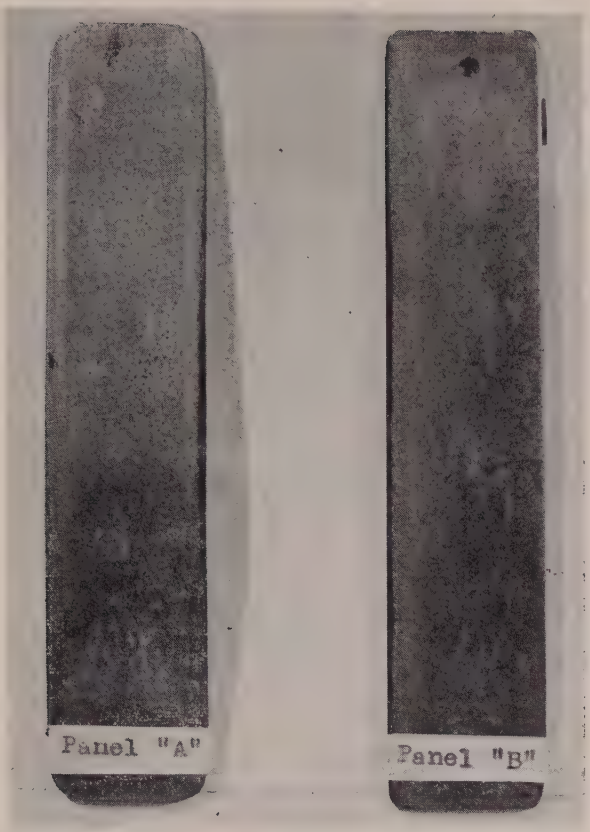


Fig. 29 — Panel A — 0.5% titanium-molybdenum coated with Chromalloy W-2 coating after testing; exposure of 2740 F for 3 min, 3420 F for 72 sec. Panel B was not tested

in the coating and no weight loss.

Recently, other cementation-type coatings have become available and are being laboratory tested at the present time with promising results.

Climax Molybdenum Co. Al-Cr-Si Coating (Flame Sprayed)

With the exception of the cementation-type coatings, this is the best coating for the protection of molybdenum against oxidation that has been tested. The coating is applied by flame spraying an Al-Cr-Si powdered mixture, which after application is diffused into the base metal at about 2000 F in an argon or hydrogen atmosphere. Samples with this coating have been tested in the thermal drop rig at temperatures up to 2800 F and were able to withstand at least six 5-min heating and cooling cycles from 2800 F to room temperature without failure. A segmented molybdenum shroud coated with the Al-Cr-Si coating has been successfully tested.

However, this coating has several drawbacks, chief among which are: that it has poor impact resistance, complex shapes cannot be readily coated, the coating builds up on the surface to several thousandths, and the surface finish is poor.

Many other coatings have been tested, but none performed well enough to be considered for applications. At lower temperatures (1800–2200 F) metal claddings and platings have shown considerable promise. However, for large sheet metal components operating at temperatures in excess of 2400 F, the cementation-type coatings show the most promise and the flame spray coatings are second. Platinum plating is a possiblbe exception to this general statement.

Conclusions

The purpose of this paper is to accurately reflect the present state of the art in the utilization of refractory metals for air frame and powerplant sheet metal components. Both the bright and dismal aspects of the use of refractory metals is presented in as unbiased a manner as is possible. It is the hope of the authors that the faults of refractory metals discussed herein will not be a source of discouragement to the development and utilization of these materials, but rather will be a challenge to the metals industry to overcome the shortcomings and provide the most efficient materials systems possible for + 2000 F structural applications.

DISCUSSION

Fabricating Failures with Molybdenum Are Mechanical in Nature

— W. L. Bruckart
Universal-Cyclops Steel Corp.

MANY OF THE points which are raised in this paper concerning the weaknesses of molybdenum are currently being worked on with the idea that a solution in a reasonable time can be expected. Property variations, high transition temperature, sheet surface contamination, and similar problems may very well prove to be a thing of the past when our In-Fab operation becomes a reality some time within the next 12 months. Further, new and better alloys of the molybdenum-base type may very well be associated on a concurrent basis with the advent of In-Fab.

I should like to take some technical difference with the authors with regard to their comments on laminations. In the first place, laminations as defined by metallurgists generally refer to claims of inhomogeneity, compositionwise, such as might be the case where nonmetallics are strung out in a wrought material. These are necessarily planes of weakness and are certainly contributing to early fracture in many cases. Such, however, is not the case in molybdenum. It should be understood that these comments from the authors and their illustrations, are observations and discussions of cracks that have been found in the metal, either as it has been worked, or as it has been received.

There are only two types of these cracks to my knowledge. One type, a fracture from manufacturing, is evidenced by the discoloration of the interface at the crack. Such a discoloration denotes that the material had been cracked at some early stage of manufacture and that the surfaces had been given the opportunity to oxidize. This is not a lamination, but a defect which was not properly inspected or observed prior to shipment.

The second type of crack, is what is considered to be a secondary tensile failure. For a detailed discussion of this type of crack I would refer readers to "Production and Quality of Molybdenum Mill Products"[a] and to a technical paper.[b] In brief, the existence of these cracks is the result of tensile stresses developed in the forming of the material, either too cold, or by improper use of tools.

There are ways of minimizing, or even eliminating these

[a] "Production and Quality of Molybdenum Mill Products," by W. L. Bruckart, in "Fabrication of Molybdenum." Pub. by ASM, 1959.
[b] "Fabrication of Refractory Metals," by W. L. Bruckart. Paper No. 177, American Society of Tool Engineers.

cracks. One such way would be to not provide the extreme amount of hot-cold working to the metal such as has been the practice in the past. Rather, apply a lesser degree of final hot-cold work, and accept somewhat of a higher ductile-brittle transition temperature which might well be associated with grains having lesser deformation. The net result of the lesser amount of final hot-cold work is the elimination, or minimization, of the tendency to provide a micacious grain structure form (which has been associated with extreme amounts of hot-cold work in all metals). Universal-Cyclops Steel Corp. is currently practicing this, and has found that the upward shift in the ductile-brittle transition temperature has not been of any consequence, but that the tendency to form these secondary tensile failures as a result of bending, sawing, shearing, and the like, has been reduced to a very great degree and shows promise of being eliminated.

In short, the problems associated with failure to fabricate in molybdenum are problems which appear to be mechanical, to a great extent, and maybe purely mechanical rather than metallurgical. It is anticipated that with the increase in working temperature, the increase in the purity of a nonreactive atmosphere, and in the general improvements in chemistry of the starting stock that molybdenum will find a greater number of uses, and a considerably brighter future through the elimination of the difficulties we are now experiencing during this pioneering stage of development.

Rolling of Molybdenum-Titanium Alloy A Difficult Production Problem

— Norman L. Deuble
Climax Molybdenum Co.

THE READING of this paper has been most gratifying. Since I have been in personal contact with the authors, I am familiar with the outstanding job they have accomplished in fabricating molybdenum and molybdenum-titanium allow sheets. Since they have extolled into advantages of molybdenum it is not necessary for me to add to their comments. However, some remarks in regard to the shortcomings which they have mentioned are in order. I would like to describe briefly what the Climax Molybdenum Co. is doing to overcome some of these difficulties:

Hot-Working of Ingots

If we were to try to forge, or roll, an ingot of one of the "super" alloys at 1200 F, most of you would say that we were not using good judgment. Yet today, we are breaking down molybdenum ingots by extruding them at less than one-half of their melting point. Plans are now being made to present to the Air Force a modification to the press at Canton Drop Forging & Mfg. Co. (they are extruding practically all of the molybdenum castings) to increase its speed, and to add heating equipment so that molybdenum, molybdenum-base alloys and other refractory metals can be heated to temperatures in excess of 3500 F for extruding. This will give us true hot working, which is a step towards eliminating some of the difficulties which have been encountered. It is quite possible that starting with a hot-worked sheet bar we will lessen the tendency towards transverse weakness and 45-deg brittleness in sheets.

Fundamental Research

We heartily agree with the authors, that much work in this field is necessary. By changes in the handling of powder we have lowered many of the impurities; however, this has shown some improvement but it has not brought about all the ductility which we think will ultimately be possible. The determination of oxygen, as mentioned in this paper, is very difficult, and varies between laboratories. We are now working on a new method to determine oxygen, with the hope that consistent results can be obtained.

By lowering the impurities and interstitials, we may lower the bend transition temperature. While on this subject, it would appear to me that a standard test needs to be developed, either the bend test or impact test, which would evaluate sheets before further fabrication. I do not know of a standard test of this type. In our work, we are using a fixture which bends the material approximately 140 deg over a 1 T radius. The strain rate which we use is 19 in. per min. This is a rather severe test compared to other bend tests with which we are familiar. With our bend tests, we have produced some sheets with a transition temperature well below 70 F.

Laminations (so-called)

Mr. Bruckart has covered this item quite well in his discussion; however, I believe that hot working up to the sheet bar and then properly reducing the sheet bar to sheet, will increase the strength through the thickness of the sheet and lessen the weakness which shows up on further fabrication.

Uniformity of Product

Since the mechanical properties in molybdenum are developed by cold working (this may be as high as 2200 F), it is necessary that the rolling temperature and the reductions be carefully controlled in order to give uniform properties throughout the sheet. While this sounds easy, it is quite a problem with molybdenum which has a high thermal conductivity and a low specific heat, with the result that after one or two passes the ends are considerably colder than the center. Special rolling techniques will be necessary in order to produce an absolutely uniform product. We are glad to learn of the improvements reported in quality of sheets accomplished in the past year which the authors mention.

45-deg Cracking

This condition probably is also tied up with the processing of the sheet and I believe will be improved when we are able to hot work the sheet bar prior to reducing it to the final sheet size. This difficulty is more prevalent in large sheets.

Uniformity of Recrystallization

As you know, the recrystallization temperature depends upon the amount of strain hardening and is a time temperature phenomena. As processing of sheets is improved, we expect a greater uniformity in the recrystallization temperature in sheets of the same gage.

In conclusion, may I again congratulate the authors on the fine work they have done on molybdenum metal? I might state that the Climax Molybdenum Co. dropped out of the sheet business some time ago because of the difficulties pointed out in this paper. We have started a research program on the rolling of the molybdenum-titanium alloy sheets, taking into consideration various reductions, recrystallization, and stress-relieving temperatures. At the conclusion of this program, we believe that we will know more about the rolling of molybdenum-titanium alloy sheets.

In addition, we have another program where we are exploring the effect of ingot quality on the properties of the final product. In this program we are making one heat of each analysis from the same powder mixture, in order to determine if the powder is a big factor. This work is being done on five lots of powder so that we should have a definite trend in regard to the effect of ingot quality.

We hope that this program will solve some of the problems which were brought up by the authors and, if we are successful, we will again be producing commercial sheets.

Authors' Closure To Discussion

ACTUALLY, the only difference that exists between Mr. Bruckart and the authors is word definition. We have chosen to call sheet separations, planes of weakness, micaceous behavior, et al., by the general term, "laminations." Mr. Bruckart does not like this word, lamination, and prefers to call the defects "planes of weakness." Whatever the terminology we all agree that the defects exist and also agree on their causes.

Road Performance of a Comprex Supercharged Diesel Truck

Max Berchtold and H. P. Gull
I-T-E Circuit Breaker Co.

THE COMPREX SUPERCHARGER has been installed and road tested in a White truck tractor (Model 9000) and with a Cummins diesel engine (JT-6-B-20°). This paper reports the results of the tests.

The Comprex delivers high air density over a wide engine speed range, and responds immediately to load changes. The authors report that the supercharge gives desirable operating characteristics for a truck diesel engine.

Response, clean exhaust, fewer gearshifts, and improved fuel consumption are among the advantages the authors describe.*

ONLY IN RECENT years has turbocharging been introduced to vehicle diesel engines. The wide speed and load range requirements, the need for ruggedness and instant response, as well as the relatively small airflows, have offered a challenging problem. The Comprex,[1] employing a new concept, can meet these requirements.

History

Over 30 years ago, Dr. Moss in this country and Dr. Buchi in Switzerland recognized the potential of turbosupercharging and introduced it successfully on ships, locomotives, and aircraft engines. The further development of the turbocharger led to the development of the gas turbine.

Ever since the early days of heat-engine development, it has been the dream of engineers and inventors to perform the always necessary transmission of energy from the expanding gas to the gas to be compressed in a direct exchange. Patents for devices to serve this purpose date back as early as 1913. Claude Sieppel of the Brown Boveri Co. was the first to propose a device with the consistent utilization of nonsteady flow phenomena. The machine performed rather well; however, a great effort had to be made to find a simple and reliable design. The Comprex, designed and adapted to diesel supercharging by the Special Products Division of our company, works essentially on the same principle as the Brown Boveri device. The direct transmission of engine exhaust gas energy to precompress the engine intake air is accomplished by means of compression and expansion waves. This aerodynamic principle of nonsteady flow in comparison to steady flow has an analogue in the alternate electric current (a-c) as opposed to direct current (d-c).

The preliminary investigation and correlation of theoretical analysis with performance tests was sponsored by the Bureau of Aeronautics, U. S. Navy Department. The development of the supercharger was initiated by the Bureau of Ships. The development of two different prototypes was carried out as a company-financed project starting in 1954.

In a cooperative program with the Caterpillar

* Paper presented at SAE National Diesel Engine Meeting, Chicago, Oct. 27, 1959.

[1] "Comprex" is a trademark of I-T-E Circuit Breaker Co.

Tractor Co. and the White Motor Co. considerable laboratory and field experience has been gathered. The practical merits of the device have been established and many improvements have been accomplished as a result of this work. In this paper, the particular road performance with a Comprex supercharged White truck Model 9000 with a Cummins JT-6 engine is being reported. The characteristics can be summarized:

1. The Comprex delivers high air density over a wide engine speed range.
2. The Comprex responds immediately to load changes.
3. The Comprex is a simple, rugged, low-speed device.
4. Twenty thousand miles of road testing have been accumulated.
5. Altitude tests up to 12,000 ft have proved the feasibility of the device to maintain full power at altitude.

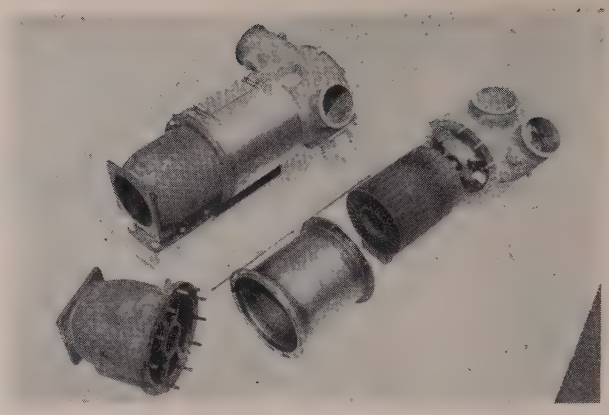


Fig. 1 — Comprex components

Principle of Operation

Idealized Cycle — The Comprex has one rotating element in which both compression and expansion take place. The four major components of the machine are shown in Fig. 1. The expanding gas acts directly, that is, in direct physical contact, on the air to be compressed. This action occurs within the straight passages of the rotor. The rotation serves only as a means of connecting each channel with the *desired* port at the *desired* time for the *desired* duration.

In the Comprex, consecutive compression and expansion waves traverse the channels and thereby transmit energy from one gas to another. In the design of reciprocating-engine manifolds, considerable experimental efforts have been made to utilize such waves to improve the filling of the cylinder. These mostly experimental efforts brought only modest returns. In the Comprex, the theory of nonsteady flow permits an accurate analysis of a cycle. It is possible to optimize such parameters as geometrical porting, rotor rpm, rotor size, air and gas intake temperatures, intake and discharge pressures, air and gas flows. It is due to this method that much progress was made in this field.[2]

The actual operation can be made clear if the sequence of momentary conditions are shown as they exist in the rotor channels. Fig. 2 shows the straight cells of the rotor and the duct connections in an isometric view. Fig. 3 shows the rotor channels unwrapped with the necessary connections to and from a diesel-engine cylinder. The aerodynamic status (location of wave front and interface, flow direction, and pressure) is indicated throughout the cycle. The port openings and duct connections to and from the engine intake and exhaust manifolds show the external flow of high-pressure air and high-pressure exhaust gas.

The turning of the rotor is equivalent to the consecutive observation of the channels from the top down. Each channel passing over the various ports is opened and closed off in the desired sequence. The cycle can best be explained if the sequence of conditions starts with a channel filled with air at rest at near-ambient pressure. The channel is closed on both ends. As the channel moves down it opens first at the port (exhaust gas intake which receives the exhaust gas from the engine. Due to the restriction this opening offers to the continuous flow of gas, the exhaust manifold pressure is considerably higher than the air pressure in the channel. When exhaust gas starts to flow into a particular channel, it has to push the air ahead of itself. This disturbance created by the hot gas entering the channel travels through the channel at a speed faster than the velocity of sound. Ahead of the disturbance (shock wave) the air is still at rest at the original pressure. Behind the shock wave, the velocity of the air equals the velocity of the exhaust gas in the channel. The static pressures in the air and the gas are equal. As this shock wave (compression wave with steep front) travels through the entire channel, the rotor turns further. The port connected to the engine intake manifold opens when the wave arrives. The compressed air will flow from the channel into the manifold. The head of the hot gas front (interface) remains considerably behind the shock wave. By the time this interface nears the channel end, the air discharge port is being closed by the further rotation of the channel. If the gas column would be in full forward motion by this time, a second compression wave would be generated by the instant deceleration at the now closed end of the channel. However, the closing edge of the exhaust gas intake port stops the flow and thus generates an expansion wave. This closing edge is so located that the entire exhaust gas column has come to rest when all the air contained in the channel has been discharged. This also is the instant when the discharge port closes. The expansion wave has decelerated the moving gas to no flow. The pressure, which has been lowered by expansion, however, is still higher than ambient pressure. Therefore, at the opening of the low-pressure discharge port, the exhaust gas starts to flow out. Again, it takes time for this acceleration expansion wave to travel through the channel. When the entire column of exhaust gas is in motion, the rotor has turned to a point where the intake port opens, and it is the inertia of the gas column which draws a fresh air charge into the channel. The intake and exhaust ports have to stay open for a sufficient time to complete this low-pressure scavenging. Then the intake port is closed first. Due to the remaining flow velocity, a third

[2] "Wave Diagrams for Nonsteady Flow in Ducts," by George Rudinger. Pub. by D. Van Nostrand Co., Inc., New York.

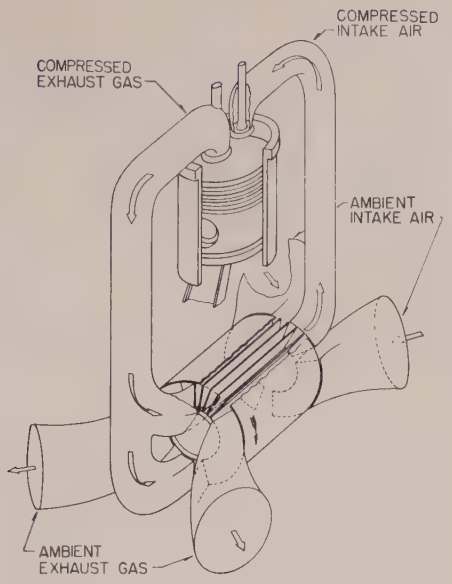


Fig. 2 — Rotor and ducting arrangement

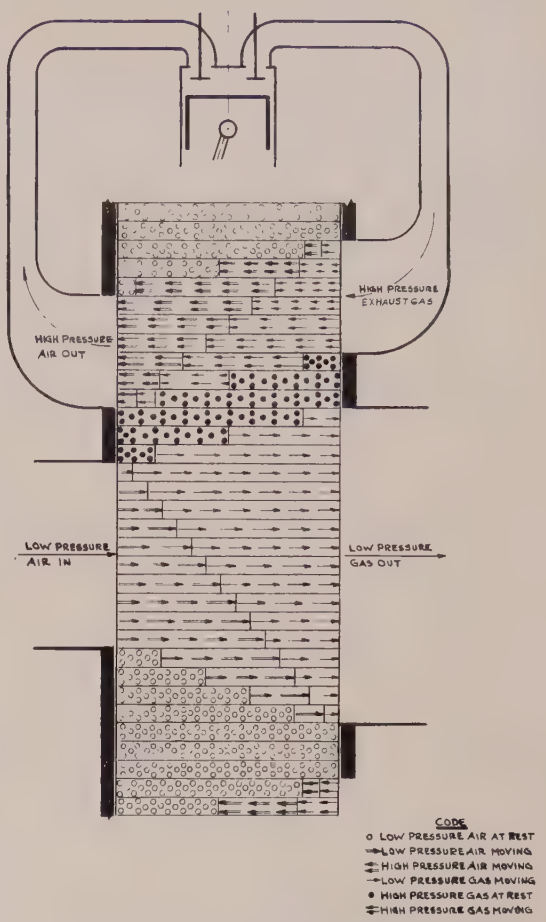


Fig. 3 — Unwrapped rotor channels wave flow diagram

expansion wave (decelerating expansion wave) originates and expands the fresh charge below ambient pressure. The exhaust port is closed at the time this expansion wave has traveled through the rotor. The low pressure now in the channel has been found desirable inasmuch as the exhaust gas velocity entering the channel is higher due to a higher pressure differential. This particular cycle has several distinct advantages for the supercharger application.

In the present arrangement, the stators are so designed that a channel goes through two complete cycles during one revolution. This design offers the advantage of symmetry and more compactness.

Deviation from Idealized Cycle — In reality there are not only the few discrete waves described. Due to the finite opening time of the channels, the waves are not sharply defined. There are several partially emitted waves which are partially reflected near the air discharge port. Similar conditions exist at the closing of the exhaust gas intake port, as well as on the low-pressure ports. Furthermore, mixing takes place at the interface between the air and the gas, resulting in an air temperature higher than the isentropic compression temperature. This effect can be minimized by a timing correction. The closing edge of the air discharge port is moved so that the contaminated air remains in the rotor and is discharged into the exhaust.

The rotor, being exposed to both the air and the exhaust gas, assumes a temperature in between the air and the gas temperatures. This is an advantage with respect to rotor temperatures, inasmuch as carbon steel can be used for the rotor. It is a disadvantage because this additional heating of the air, as well as the heating by mixing, reduces the density gain which could be obtained by isentropic compression. It has to be kept in mind that this heating is not due to internal flow losses as they occur in turbocompressors. The actual compression is rather efficient. The major energy losses occur as leakage loss at the high-pressure ports. It is, therefore, essential to maintain small clearance between the rotor and the stator. The kinetic energy lost in the exhaust represents another substantial portion of the overall losses. This energy loss, however, is utilized to help overcome the pressure drop in the air cleaner, exhaust pipe, and muffler; a loss which usually creates additional engine losses. The deviations from the simple idealized cycle can all be taken into account in the analysis. The method of calculation described by Burri[3] gives a good agreement between calculated and measured performance. The method is particularly helpful in the study of relative magnitude of different losses. Furthermore, it is possible to analyze the performance at different Comprex speeds and airflows in order to optimize the operating range.

In the Comprex, several channels discharge compressed air simultaneously into the engine intake manifold, as shown in Fig. 3. There are also several channels simultaneously exposed at the exhaust gas intake port resulting in a steady intake flow. On a multicylinder diesel engine, it can be assumed that the exhaust gas manifold and the intake manifold are sufficiently large to eliminate fully the pressure fluctuations caused by the blowdown and the instant demand of each cylinder. The pulsating flow which is generated by the opening and closing of the channels is limited to the rotor channels. The basic operation of the Comprex is in no way related to the

[3] "Nonsteady Aerodynamics of Comprex Supercharger," by H. U. Burri. ASME Paper 58-GTP-15, 1958.

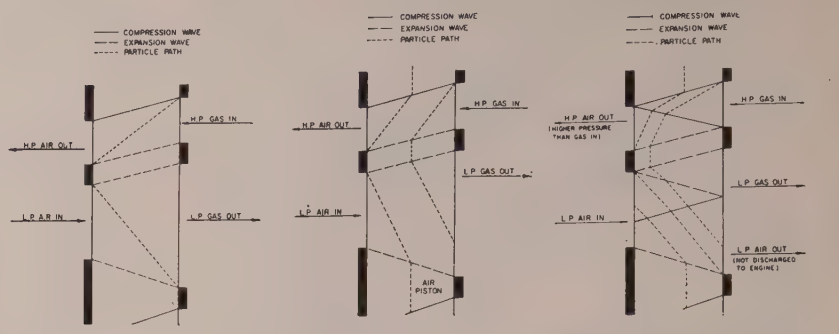

COMPRESSION WAVE
EXPANSION WAVE
PARTICLE PATH
H.P. GAS IN
H.P. AIR OUT
L.P. AIR IN
L.P. GAS OUT
AIR PISTON
H.P. AIR OUT (HIGHER PRESSURE THAN GAS IN)
L.P. AIR OUT (NOT DISCHARGED TO ENGINE)
A) DESIGN CYCLE
B) LOW PRESSURE CYCLE WITH INACTIVE GAS PLUG
C) REDUCED FLOW CYCLE WITH DECELERATION COMPRESSION WAVE

Fig. 4 — Simplified wave flow diagrams

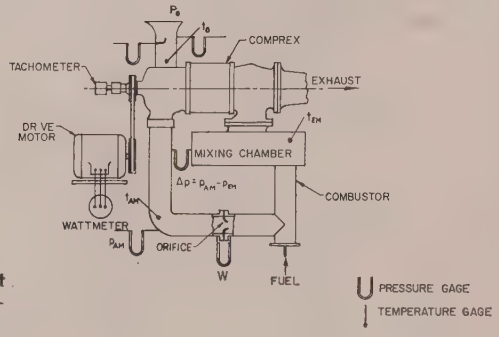


Fig. 5 — Test stand instrumentation

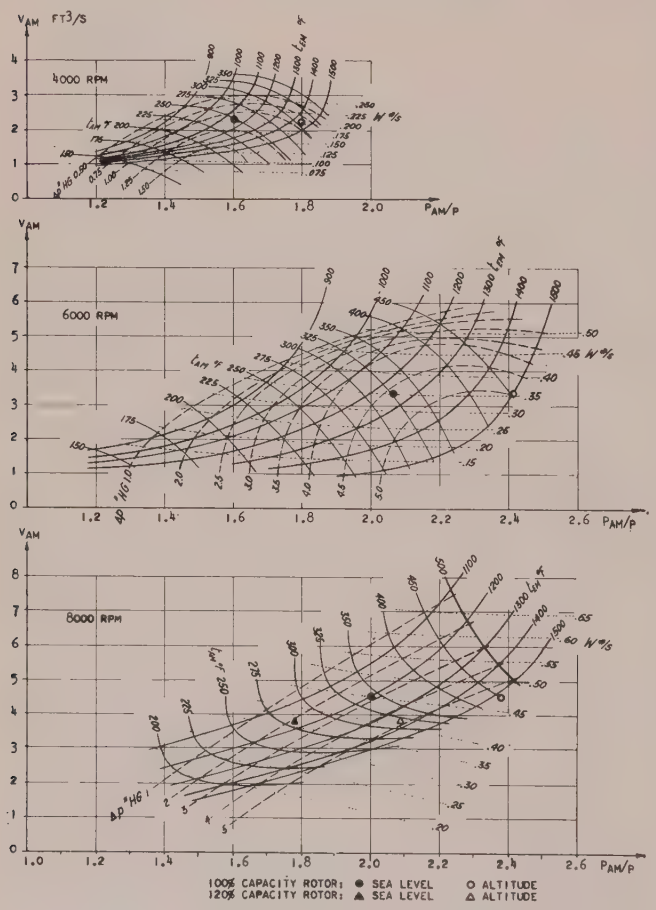


Fig. 6 — Comprex test stand performance for 4000, 6000, and 8000 rpm (ambient: $t_o = 90$ F and $P_o = 30$ in. of Hg, with selected full-load operating points)

blowdown pulses generated by the engine.

It would appear that the Comprex being so dependent upon the proper wave timing would operate efficiently only near the design point. Tests, however, have proved that this device has a very broad operating characteristic. All parameters — air and exhaust gas pressure ratio, mass flows, rotor speed — can be varied over a wide range. A few typical variations are shown in Fig. 4. Fig. 4A is a simplified presentation of the cycle diagram shown in Fig. 3. If, for instance, the manifold and exhaust pressure is reduced, the rotor flow velocities decrease. The interfaces penetrate less into the rotor. As a result, a plug of air remains in the rotor as shown in Fig. 4B. The Comprex can operate at variable pressure ratio. In Fig. 4C, a cycle is shown where the engine manifold pressure is higher than the engine exhaust pressure. A reflecting compression wave decelerates the flow and increases the pressure. Only a part of the available air in the rotor is discharged. The air remaining in the rotor expands again and is discharged with the exhaust gas. The Comprex has no surge limits; it adapts itself to a variable mass flow demand.

Diagrams for off-design speed operation become much more complicated due to wave reflections. It is, however, possible to calculate the off-speed cycles with the established calculating methods. Based on these results, modifications of the ports can be made to accomplish a specific purpose, such as a wide speed range. It is in this field of porting design compromises that considerable progress has been made over the past year.

The Comprex rotor, which performs the function of a rotating timing valve, does not produce or absorb power except for bearing friction and windage losses and, therefore, has to be driven from the engine. The simplest and most reliable drive is a direct "V" belt or gear drive which determines the Comprex rpm to be proportional to engine rpm. This, however, is not the most desirable speed schedule to obtain optimum Comprex performance. Since special drives are too unreliable and too costly to be practical, an improvement of the Comprex speed range was necessary before this simple drive arrangement could be used.

Comprex Calibration

To analyze and predict the performance of a Comprex supercharger with a specific geometrical configuration, the complete characteristic (variation of

all parameters) must be known. The characteristic has been measured on the test stand shown schematically with the instrumentation in Fig. 5. Fig. 6 shows the characteristic. The effect of the diesel engine is simulated by a combustion chamber and an orifice. The performance for a given cycle (porting geometry) is described by three individual charts, each for a different Comprex speed. Air discharge volume V_{AM} is plotted as a function of the air discharge pressure ratio P_{AM}/P_o for constant gas intake temperatures t_{EM}, constant air discharge temperatures t_{AM}, constant flows W, and constant air discharge to gas intake pressure drops ΔP. The plot is made for $t_o = 90$ F ambient temperature and $P_o = 30$ in. of Hg. For a different ambient temperature t_o', the correction factor to be applied is:

$$\theta = \frac{t_o' + 460}{t_o + 460} \tag{1}$$

For different ambient pressures P_o', the correction factor to be applied is:

$$\delta = \frac{P_o'}{P_o} \tag{2}$$

The volume flow V_{AM}' corrected to the new ambient condition P_o' and t_o' is then:

$$V_{AM}' = V_{AM}\sqrt{\theta} \tag{3}$$

The weight flow W' corrected to the new ambient condition is:

$$W' = \frac{W\delta}{\sqrt{\theta}} \tag{4}$$

The pressure drop $\Delta P'$ for the new ambient condition is:

$$\Delta P' = \Delta P\delta \tag{5}$$

If the rotor size is changed from the present rotor flow area A to a new rotor flow area A', the corrected volume flow is:

$$V_{AM}' = V_{AM}\frac{A'}{A} \tag{6}$$

and the corrected weight flow is:

$$W' = W\frac{A'}{A} \tag{7}$$

If the rotor diameter is changed to fulfill the capacity requirement, it is also necessary to change the rotor rpm and the rotor length. In order to maintain the same Comprex performance shown in Fig. 6, all absolute and relative flow velocities have to remain constant at a certain operating point. This means the tip speed has to be the same; therefore, the new rotor rpm will be:

$$n' = n\frac{d}{d'} \tag{8}$$

and the new length to retain the original wave timing will be:

$$l' = l\frac{d'}{d} \tag{9}$$

where:

d = Mean rotor diameter of rotor tested
d' = new mean diameter

The rotor dimensions of the calibrated unit are:

Hub diameter	OD	2.875 in.
Shroud diameter	LD	5.55 in.
Mean diameter		4.21 in.
Flow area		14.8 in.^{2}

Design Description

Fig. 7 shows two longitudinal sections through the Comprex with both the low-pressure and the high-pressure flow passages in both stators. The porting can be seen in the cross-section. The air enters the intake stator through two independent passages to the two intake ports. The compressed air discharge ports lead to a common collector which makes possible the use of a single connection to the engine. An alternate stator design has been completed with a common air intake instead of the two separate intakes. This will make the installation on the en-

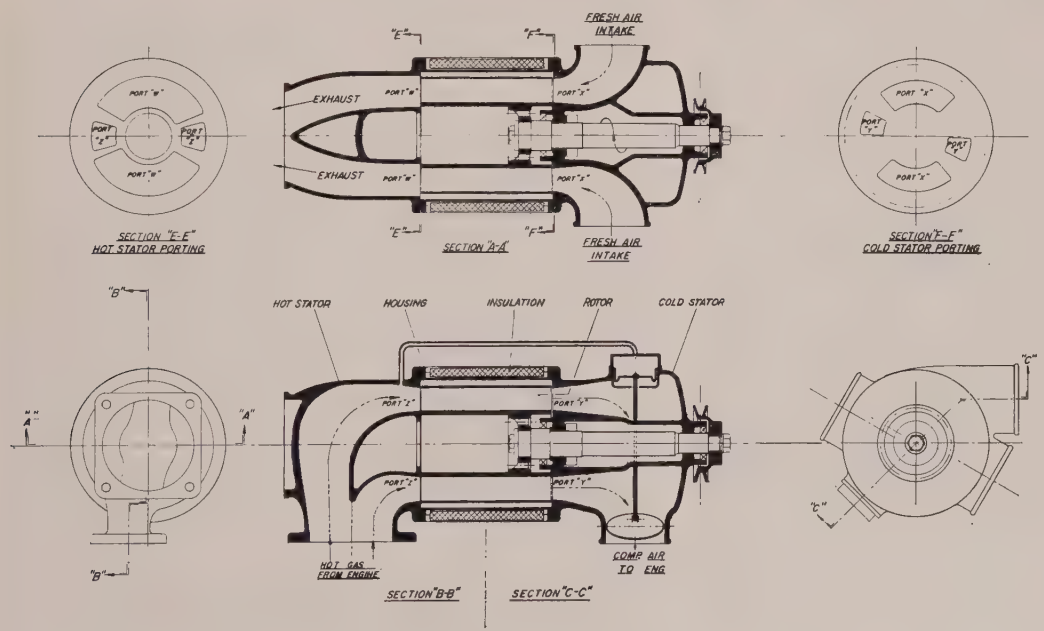


Fig. 7 — Comprex supercharger cross-section

gine more compact.

The two sleeve bearings for the cantilever rotor shaft are located in the air intake stator (cold stator). The bearings being housed in the same casting eliminate problems of alignment. With the low bearing load and the low shaft speed of 9,000 rpm, the use of aluminum sleeve bearings with normal clearances is permitted. The bearings are lubricated with engine oil. No special filtration of the oil is required.

The alloy for the cold stator is aluminum 356-T7. The cast aluminum is used directly as bearing surfaces. The intake stator casting which houses the airflow passages, the bearings, oil seals, and oil sump also contains the throttle and the throttle actuator required for starting and idling. In addition, the modified intake stator will house the bypass valve also necessary for starting.

The hot stator has a common exhaust gas intake flange. Two separate passages lead to the two high-pressure ports. The two low-pressure ports connect to a single exhaust flange. The hot stator alloy is NiResist 2B. This material gives excellent dimensional stability without cracking or flaking after continuous exposure to the variable exhaust gas temperature.

The rotor, which represents the most important component, has to operate under the most adverse conditions. It has a cylindrical shroud and a cylindrical hub with 35 straight axial passages of constant flow area. At the air intake side, the rotor assumes about air discharge temperature. At the hot end, the rotor temperature stays considerably below the exhaust gas temperature due to the extra blowthrough of the air during the low-pressure scavenging. Nevertheless, the temperature gradient along the axis is considerable, but fortunately it is about linear. The cylindrical hub, therefore, assumes a cone shape. The blades, held by brazing in straight parallel or dovetail grooves, do not restrain the hub from this deformation — provided the expansion of the hub is unrestrained by the shroud. It is, therefore, necessary to provide expansion joints in the shroud. The shroud is formed of five overlapping segments. The method of attachment of the blades to the shroud segments is the same as on the hub. Each segment has seven blades. Comparative tests with 49, 35, and 27 blade rotors were made to establish the optimum blade number.

This rotor design has been tested over 1000 hr with hot gas temperature cycling. Rotors with incomplete brazing failed after 50 hr. With the proper control of the groove and blade dimensions and the brazing cycle, the required quality of the joints for the required rotor life is obtained. The stresses due to centrifugal loads are low on account of the modest rotational speed (9000 rpm with a 5.5-in. rotor diameter). In an alternate rotor design, a continuous shroud is used in connection with nonradial blades, which allow deflection in radial direction to accommodate differential expansion of hub and shroud. With this simple rotor design (straight grooves in hub and shroud), it is possible to reduce the required driving power by a slight local bend of the blades near the hot end of the rotor. The drive power is approximately 1% of the engine output power.

The thrust bearing which controls the clearance between the rotor and the cold stator is located near the stator face. This arrangement assures approximately constant clearance between the cold stator and the rotor regardless of temperatures. The clearance between the rotor and the exhaust gas stator (hot stator) is controlled by the expansion of the rotor housing. The rotor housing is insulated at the outside. The distance between the rotor shroud and housing is held at close proximity in order to assure an effective heat exchange between the rotor to the rotor housing. The compressed air leaking into the annular chamber between the rotor and the rotor housing is drained into the exhaust port, utilizing the relative vacuum created by the high velocity exhaust gas. The wall thickness of the housing is small to reduce the lag between the rotor and housing expansion to a minimum. The tests have shown that this simple system is capable of maintaining the hot end clearances sufficiently constant even during transient conditions. The cold end assembly clearances are held between 0.008 and 0.010 in. and the hot end assembly clearances between 0.014 and 0.016 in.

The hot stator serves as the main support for the unit. This system, which eliminates the need for a hot expansion joint, has been used on most small turbochargers. The Comprex being somewhat smaller in diameter but longer is further supported by a flexible strut near the cold end. During road tests this system has been found fully satisfactory.

The weight of the Comprex is about the same as the Roots blower or the turbocharger of equal capacity. The design of the Comprex in its present form is extremely simple. The intricate flow passages can easily be produced in the castings. The rotor design lends itself to simple manufacturing techniques. With the carbon steel rotor, the need for alloy steels is reduced to a minimum. The Comprex represents, therefore, a simple low-cost device.

Vehicle Power Requirement

In order to discuss the problems of matching the Comprex to a vehicular engine, it is important to review the requirements.

The ideal powerplant would be an engine capable of constant power independent of engine speed. This requirement has been set forth in an analysis of future engines by Marks and Reiners.[4] This would mean a torque inversely proportional to engine speed. A conventional engine with an infinitely variable gear ratio which can be shifted under power would be essentially equivalent since the engine would operate at maximum engine speed regardless of vehicle speed. However, gear shifting under load needs complicated mechanisms. Both versions are remote to practical realization. The practical solution lying between these extreme cases is acceptable if constant power can be maintained over a limited engine speed range, so that with a finite number of gear ratios the power output remains constant over the entire vehicle speed range.

The power required by a vehicle is composed of the friction power caused by air drag and tire friction and the power to lift the weight (gross vehicle

[4] "Forward Look on Engines," by D. T. Marks and N. M. Reiners. Paper presented at SAE National West Coast Meeting, Los Angeles, August, 1958.

[5] "Factors Affecting Fuel Economy (MPG)," by J. C. Miller. Paper presented at SAE Summer Meeting, Atlantic City, June, 1957.

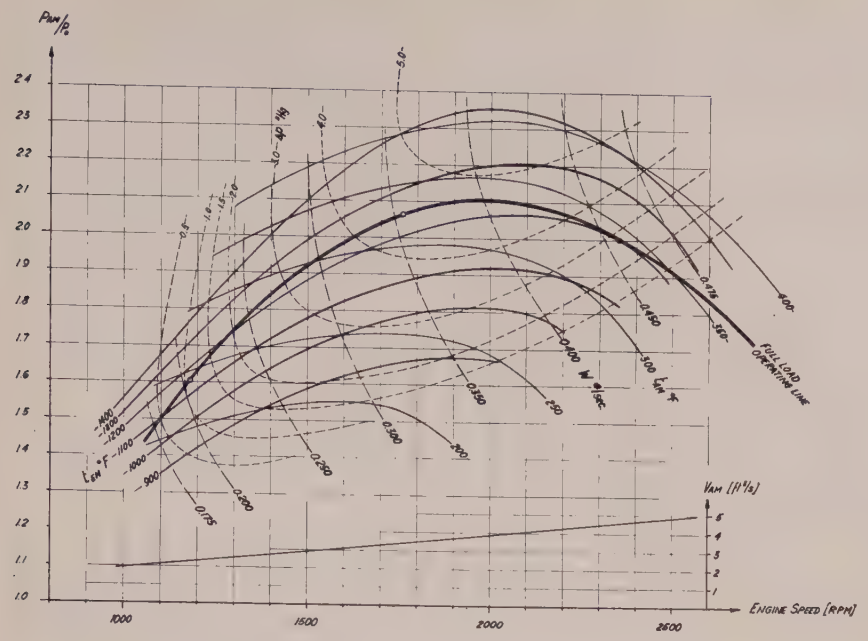


Fig. 8 — Engine performance predicted from test stand performance (ambient: $t_0 = 90$ F and $P_0 = 30$ in. of Hg, with full-load engine operating line, 100% capacity rotor at sea level)

weight). In Miller,[5] the friction power for zero wind, zero grade for 57,500-lb gvw is given as a function of vehicle speed. Since the tire friction power is proportional to the square of the vehicle speed, and the wind friction power is proportional to the third power of the vehicle speed, the combined friction power is found, as expected, to be:

$$N = C_D V^{2.25} \tag{10}$$

where:

N = Horsepower to maintain vehicle speed, V
C_D = Combined drag coefficient (air drag and tire friction)

The drag force F_D is:

$$F_D = C_D V^{1.25} \tag{11}$$

The required horsepower for different per cent grades as a function of vehicle speed V and gross vehicle weight is then:

$$N = C_D V^{2.25} + (\text{gvw})\ V \sin \alpha \tag{12}$$

where:

α = Angle of the grade incline

A truck with a given maximum road horsepower and the ideal transmission will travel at each grade with the maximum speed.

Matching Comprex to 4-Stroke Vehicle Diesel Engine

In order to obtain an engine characteristic approximating the ideal vehicular requirements as outlined above, it is necessary to schedule the fuel as well as the airflow. If the available exhaust energy can be utilized to furnish the required airflow at all engine operating points with a simple mechanical device, the supercharged diesel engine will represent the most desirable powerplant from the standpoint of engine flexibility, fuel consumption, and powerplant weight. The Comprex properly selected can fulfill these conditions.

The matching procedure represents a problem of successive approximation. The solution obtained for the actual case of the truck tests at sea level and altitude is given, and desirable improvements are outlined.

The naturally aspirated engine operates essentially at a constant air filling of the cylinders independent of engine speed and load. At full load the exhaust temperature reaches a maximum at an air fuel ratio of approximately 17: 1. More fuel input will only produce smoke. The engine torque is constant. The usual torque rise of about 10% is due to reduced friction power and better volumetric efficiency at lower speed. If an exhaust driven supercharger is used, the increasing exhaust temperature with increasing load furnishes the energy to obtain higher intake manifold density, which means a larger cylinder filling at higher load. The airflow at a given engine speed is, therefore, variable with load. In the vehicular engine where it is desirable to maintain the power output with reducing engine speed, it is necessary to increase the manifold density with decreasing engine speed.

In order to select Comprex parameters, the complete performance curves such as shown in Fig. 6 pertaining to a specific porting geometry are necessary. The characteristic represents the results of considerable development effort towards the most desirable matching. Further refinement is now possible with the dynamometer and road experience recently accumulated.

As previously mentioned, the Comprex characteristic Fig. 6 gives the air discharge volume flow V_{AM} as a function of pressure ratio for different exhaust gas temperatures t_{EM} as well as the values for airflow W, engine pressure drop ΔP and air discharge temperatures t_{AM}. Each of the three plots is for a constant Comprex speed.

In a first approximation, the volume demand of the engine is a function of the engine speed only. The volumetric efficiency η_v of the engine is, accord-

ing to Mitchell,[6] a function of the pressure drop ΔP through the engine. The volumetric efficiency $\eta_v = \frac{\text{intake air volume}}{\text{intake displacement}} = 1$ has been found to give good agreement at all speeds. The effect of the larger ΔP at higher speeds is compensated by the higher flow losses in the valves.

The results shown for the truck engine used during the road evaluation represent a solution found by successive approximation.

Sea Level Performance

The Comprex drive ratio, which is 3.4: 1, assigns the map Fig. 6 for 8000-rpm Comprex speed to an engine speed of 2350 rpm. The required volume flow $V_{AM} = 4.55$ ft³/sec. The desired maximum torque of the engine at this speed expressed in brake mean effective pressure (bmep) is 138 psi. The exhaust temperature for this bmep established by dynamometer test is approximately 1200 F. The plot in Fig. 6 for 8000 rpm shows an operating point of 2-atm manifold pressure, an air discharge temperature of 340 F, a pressure drop of 2.8 in. of Hg between the intake and the exhaust manifold, an airflow of 0.455 lb/sec, and an air density of 0.100 lb/ft³.

At 6000 rpm Comprex speed (Fig. 6) the engine speed is 1760 rpm; required volume flow is 3.42 ft³/sec. The desired bmep of 166 psi will require an exhaust temperature of 1250 F. This will give a manifold pressure of 2.07 atm, a manifold temperature of 325 F, an airflow of 0.355 lb/sec, and a manifold density of 0.105 lb/ft³.

At 1170 rpm the volume flow requirement of the engine is 2.28 ft³/sec; the Comprex speed is 4000 rpm (Fig. 6). The desired bmep of 153 psi requires an exhaust temperature of 1150 F, the manifold pressure is 1.60 atm, and the manifold temperature 250 F. The airflow is 0.205 lb/sec, density is 0.900 lb/ft³.

From the three Comprex performance charts (Fig. 6), the complete engine part load performance map (Fig. 8) can be extracted for the selected drive ratio. The air manifold pressure $\frac{P_{AM}}{P_o}$ versus engine speed is plotted for constant exhaust gas temperatures t_{EM}. Furthermore, lines for constant air manifold temperature t_{AM}, engine pressure drop ΔP, and air weight flow W are included. This diagram gives all the thermodynamic Comprex data as determined by the engine flow requirements and exhaust temperatures t_{EM}. This combined Comprex performance chart has been drawn beyond the engine limitations. The permissible maximal torque at each speed which is limited by the engine parameters is determined by the maximum injected fuel quantity. The limits are given by the maximum permissible exhaust valve and piston temperatures which are functions of the particular engine design. This limit, generally called the thermal loading, is not a clearly defined quantity. Specific attention has to be given in the case of the Comprex. The higher compressed air temperature will result in higher engine cycle temperatures throughout. The increased heat rejection will cause higher piston and valve temperatures. Comparative dynamometer tests, however, indicate that the increased heat loss is partially compensated by a more rapid combustion induced by higher air temperatures. Furthermore, the effect of the high manifold pressure has to be evaluated with respect to bearings, connecting rods, and pistons, particularly at low speed when the inertia forces are less and the connecting rod forces might become excessive.

Since the air manifold temperatures are lower at lower speed and since there is more time to remove the rejected heat per revolution by the cooling water, high torques offer no problem with respect to heat load. As pointed out, the desired torque

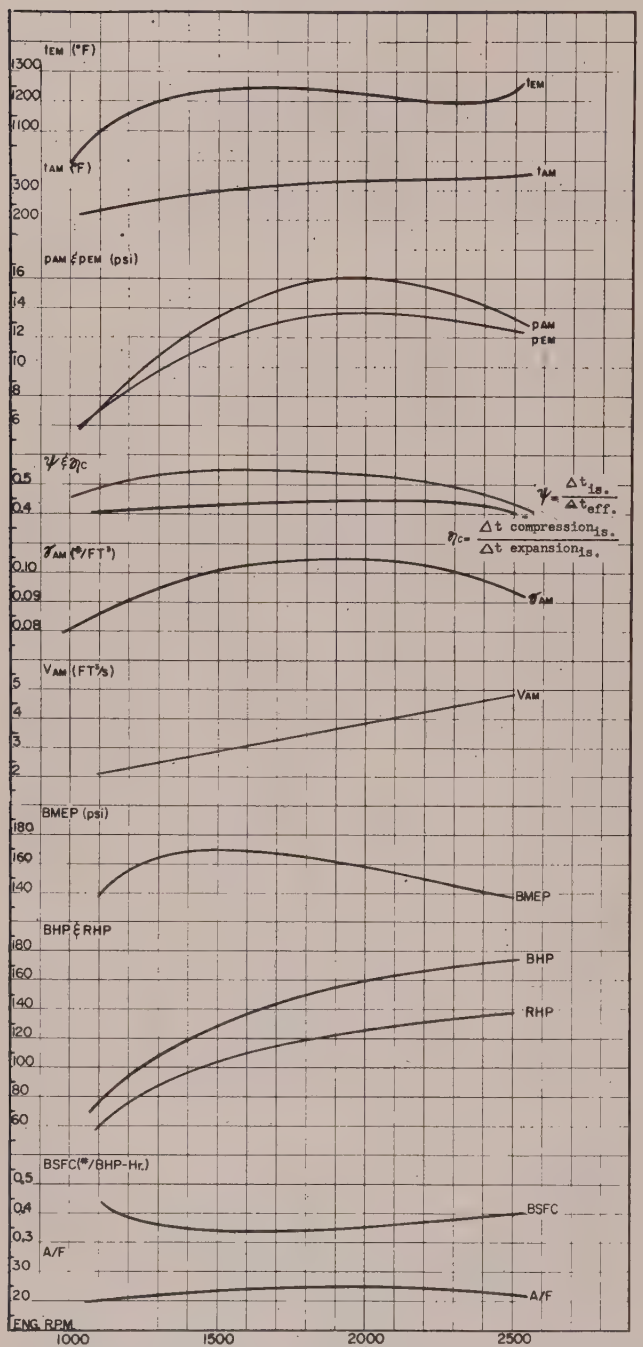


Fig. 9 — Engine performance from chassis dynamometer tests (ambient: $t_o = 90$ F and $P_o = 30$ in. of Hg, full-load engine operating line, 100% capacity rotor at sea level)

[6] "Evaluation of Aftercooling in Turbocharged Diesel Engine Performance," by J. E. Mitchell. SAE Transactions, Vol. 67, 1959, pp. 401–413.

curve within the engine limits can be obtained by programming the fuel flow as a function of engine rpm. The Cummins fuel system, for example, is easily adaptable to such requirements, as pointed out by Reiners and Schwab.[7] The operating points from Fig. 6, which represent the maximum torque determined by the maximum fuel pump setting, sets up the full-load operating line included in Fig. 8.

Fig. 9 represents a plot of the chassis dynamometer performance. As a function of engine speed, the following values are plotted: maximum road horsepower, air manifold and exhaust manifold temperatures and pressures, engine horsepower, bmep, specific fuel consumption, air fuel ratio, air manifold density, and relative temperature rise:

$$\psi = \frac{\Delta t_{is}}{\Delta t_{eff}} \qquad (13)$$

and the combined energy efficiency:

$$\eta_c = \frac{\Delta t_{\text{compression}_{is}}}{\Delta t_{\text{expansion}_{is}}} \qquad (14)$$

where

Δt_{eff} = Total work and heat input (heat transfer and mixing)

Δt_{is} = Isentropic work input

The value ψ permits a comparison with the compression efficiency in conventional compressors. From the standpoint of the engine, a high ψ is a valuable criterion, inasmuch as it is desirable to keep the air manifold temperature t_{AM} as low as possible. The combined energy efficiency η_c gives a measure of the internal losses. This combined efficiency includes the losses caused by intake throttling and muffler exhaust back pressure.

The nearly constant air density over a wide engine rpm range allows maintaining a nearly constant air/fuel ratio with the desirable torque characteristic. Together with the high response of the Comprex, it permits smokeless engine operation. This is a significant factor as smoke, in almost all cities, has become of more concern.[8]

The engine bmep shows a rise of 18% between 100 and 75% engine speed. The agreement between the predicted engine operating points and the measured engine operating point is a good indication of the validity of the matching procedure. The road horsepower values are entered in the truck speed-grade plot (Fig. 10), for the 8-speed gearbox and the rear axle used on the truck (White Model 9000). As a comparison, the same graph contains the road horsepower of a standard engine with a torque rise of 5% between 100 and 75% engine speed, using the same gearbox with an engine of equal maximum power.

The standard engine for instance has in eighth gear a maximum speed of 48 mph on the level road. The engine does not produce the maximum possible horsepower, since the only intersecting point lays at a speed lower than maximum speed. However, it is advisable to shift gears at this speed; the engine is then operating at the maximum speed and port load. For a slight increase in grade, it will not be necessary to shift gears. The speed of 48 mph can be maintained up to a grade of 0.3% where full power is needed. With an increasing grade up to 0.9%, the speed will fall to 34 mph where shifting becomes advantageous again.

The Comprex supercharged engine approaches much closer the ideal requirement of constant power. The gains appear in higher truck speed at the same grade, or a higher grade at the same speed.

For instance, in a 0.9% grade, the truck with Comprex can maintain 39 mph while the truck with standard torque curve slows down to 34 mph. In a 1.2% grade, the truck with Comprex can maintain 34 mph. This fact has proved itself in extensive road tests. The frequency of gear shifting is greatly reduced. The truck can maintain a higher average speed. In Parker,[9] the author reports on the development of a turbocharger with wastegate. This arrangement also permits an increase in torque rise and the similar observation of less frequent gear

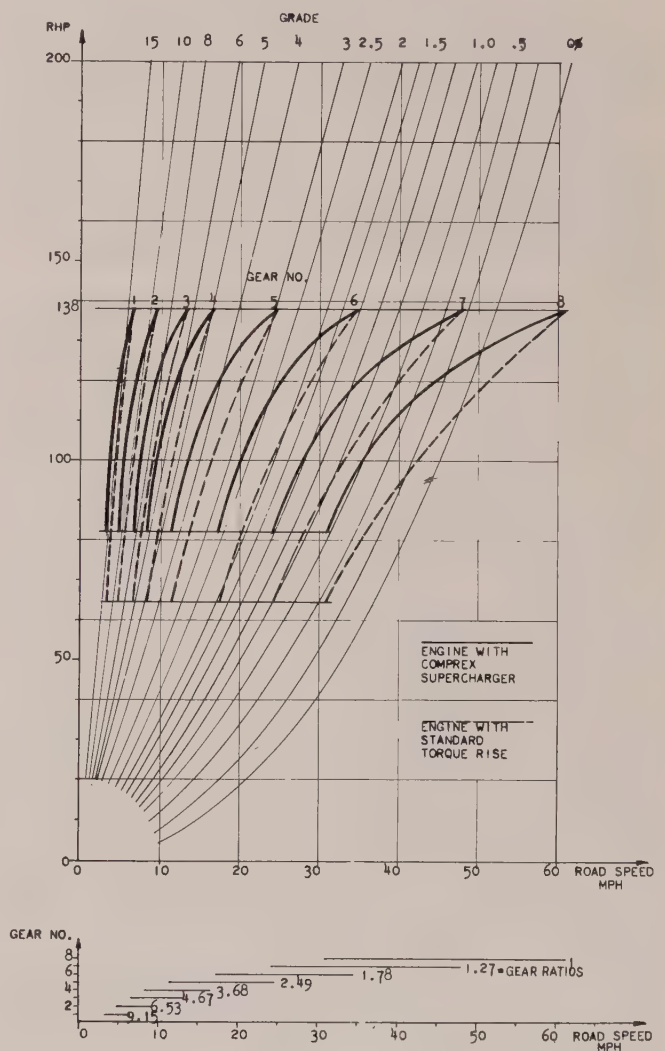


Fig. 10 — Truck power requirement for 57,500-lb gvw, variable grade in per cent, variable speed in mph, and no wind; with comparison of two engine characteristics at sea level — engine with Comprex supercharger and engine with standard torque rise

[7] "Cummins Turbocharged Diesel Engines," by N. M. Reiners and W. D. Schwab. Paper presented at SAE National Transportation Meeting, Cleveland, November, 1957.

[8] "How to Recognize Unnecessary Vehicle Exhaust Smoke Emissions," by AMA Special Group on Exhaust Smoke, F. W. Bowditch, chairman. Paper presented at SAE National West Coast Meeting, Seattle, August, 1957.

[9] "Problems in Matching Turbocharger to High Speed Diesel Engines," by Wilton E. Parker. Paper presented at SAE National West Coast Meeting, Los Angeles, August, 1958.

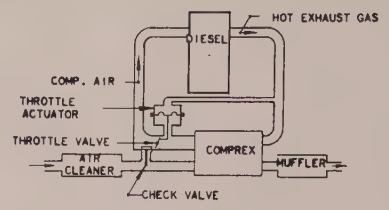


Fig. 11 — Comprex control arrangements

shifting was found to be a considerable advantage. The instantaneous response to load changes during all maneuvers of the Comprex supercharged truck has been found to be one of the most desirable features. The engine is as responsive as the much heavier naturally aspirated engine.

With the proper selection of a gearbox and rear axle, it is possible to lower the engine speed for cruising. This will give a reduced engine speed at the same truck speed and will, therefore, improve fuel economy.

Altitude Performance

To establish the altitude performance of an engine, the performance maps (Fig. 6) can be again be used. The corrections necessary to adapt the maps to altitude performance are given above.

The following evaluation shall explain the general relationship of the quantities involved. The assumptions are simplified to demonstrate the effects of major importance. The results of this analysis, though not fully agreeing with the actually measured data, nevertheless show reasonable agreement.

The altitude chosen for the analysis is 10,000 ft with an ambient temperature t_o' of 90 F. The ambient density then reduces in the ratio of ambient pressures $\frac{P_o'}{P_o} = \delta$, which for 10,000 ft is $\delta = 0.69$. With the same engine, the displacement and the rpm remain the same, and likewise the volume demand of the engine. The fuel setting being unchanged means a constant fuel input. The air/fuel ratio would decrease unless the Comprex is capable of increasing the pressure ratio to maintain the same absolute manifold pressure and temperature as at sea level. This, however, is not possible due to the pressure ratio limitations in the present Comprex cycle. The air/fuel ratio decreases, leading to a higher exhaust temperature. The higher pressure ratio and increased airflow partially compensates for the lower airflow due to altitude, thus the nozzle temperature does not increase as much. The actual equilibrium can easily be established. In the case of the 8000-rpm Comprex speed, or 2350-rpm engine speed, the pressure ratio increases from 2 to 2.38 atm and the exhaust temperature increases from $t_{EM} = 1200$ F to $t_{EM}' = 1530$ F. At sea level ($P_{AM}/P_o = 2$) the airflow was $W = 0.455$ lb/sec.

Due to the increased pressure ratio $P_{AM}'/P_o = 2.38$, the mass flow at sea level density increases to $W = 0.475$ lb/sec. At the low intake density $\gamma_o' = 69\gamma_o$ the airflow is $W_{AM}' = 0.475 \times 0.69 = 0.328$.

The rejected heat from the engine at sea level is:

$$\overset{\nearrow}{Q} = (t_{EM} - t_{AM})\, c_P W \qquad (15)$$

and at altitude:

$$\overset{\nearrow}{Q'} = (t_{EM}' - t_{AM}')\, c_P W' \qquad (16)$$

The two values can be assumed to be equal if the fuel input or added heat $\overset{\swarrow}{Q}$, as well as the engine cycle efficiency remain constant at sea level and altitude. The $(A/F)'$ is, therefore, inversely proportional to the airweight flow:

$$(A/F)' = (A/F)\ \frac{W'}{W} = 24\,\frac{0.328}{0.455} = 17.3 \qquad (17)$$

and the exhaust manifold temperature at altitude:

$$t_{EM}' = (t_{EM} - t_{AM})\,\frac{W}{W'} + t_{AM}' = 860\,\frac{0.455}{0.328} + 450 = 1640 \qquad (18)$$

The air/fuel ratio which is 24: 1 at sea level is reduced to 17: 3 at 10,000 ft. This means full power is obtained at this altitude. The computed exhaust temperature is 1640 F instead of the expected 1530 F. If $t_{EM}' - t_{AM}'$ would be 1080 F instead of 1190 F, the exhaust temperature would agree with the anticipated value (Fig. 6, 8000 rpm). This is equivalent to a 10% reduction of the exhaust heat rejection. Since the wall heat rejection increases with the higher intake temperature $t_{AM}' = 450$ F, this appears reasonable.

The measured pressure on the truck at 2300 engine rpm and 10,000 ft altitude was $P_{AM}'/P_o' = 2.38$, the observed air manifold temperature $t_{AM}' = 425$ F, and the exhaust manifold temperature $t_{EM}' = 1400$ F with an ambient temperature of about 75 F. These temperatures corrected to 90 F ambient make $t_{AM}' = 450$ F and $t_{EM}' = 1450$ F. In view of difficulties of obtaining steady-state data, this can be considered good agreement.

The altitude conditions at 1750 and 1160 engine rpm change similarly. The points computed by the same method for 6000 and 4000 rpm are also shown in Fig. 6. From the foregoing, it can be concluded that the engine can produce full power up to 10,000-ft altitude over the full engine speed range without smoking or danger of overspeed. The Comprex pressure ratio increases considerably. The flow demand of the engine, however, exceeds the desirable Comprex capacity resulting in high air temperatures t_{AM}'. The interface penetrating through the rotor causes recirculating of gas with the compressed air into the engine. The contamination and the high air temperature are both undesirable.

Improved Matching in View of Altitude Performance

With a different selection of Comprex capacity, the matching can be improved so that the air manifold temperatures are reduced, particularly at altitude. Again reference is made to Fig. 6. At lower engine rpm, the Comprex furnishes the air at lower temperature. This means that a larger Comprex can furnish the air at the desired temperature at the full engine speed. A Comprex with a 20% larger rotor flow area will mean a new volume and weight flow scale, whereby the 5 ft³/sec line becomes the 6 ft³/sec line; the 0.5 lb/sec line becomes the 0.6 lb/sec line, and the like. The same effect is obtained if instead the engine volume demand is reduced by 20%. At the 2350 rpm engine speed, the corrected volume demand becomes: V = 3.80 ft³/sec.

Assuming the fuel input and the thermal efficiency of the engine cycle remains the same, the

rejected heat will remain constant. The new operating point has again been found by successive approximation. The new exhaust temperature t_{EM} will be 1175 F; the new air discharge temperature is $t_{AM} = 280$ F and the new air manifold pressure will be 1.78 atm; the new air density 0.096 lb/ft^3. The charge density decreases 4%, which means that the air/fuel ratio decreases to 23: 1. The rejected heat is then the same as for the original size engine. With the lower air temperature t_{AM} the ψ value improves from 0.485 to 0.515. At 10,000-ft altitude the conditions are now also improved. The manifold pressure raises from 1.78 at sea level to 2.08 at 10,000 ft. The air manifold temperature for 60 F ambient will be 273 F. The gains in the air manifold temperature both at sea level as well as at 10,000 ft are quite significant, and it appears beneficial to evaluate the gains fully by test. A unit with the same outside dimensions using a rotor with 20% larger capacity is being tested.

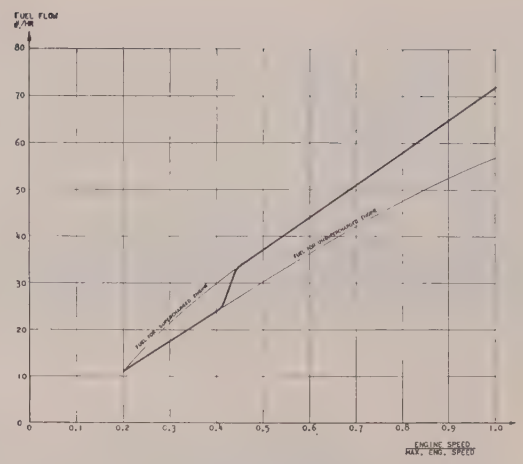


Fig. 12 — Fuel schedule for no-smoke operation

Control System

The Comprex, being a timing device, has to operate at the proper speed. With the belt or gear drive, the speed of the Comprex during cranking and idling is out of synchronization. A simple control system has been developed which throttles the airflow to the engine to such an extent that the Comprex remains operative regardless of mistiming. The delivered flow is not sufficient to meet the engine demand. A check valve within the intake manifold allows air to enter as soon as the manifold pressure drops below ambient pressure. The throttle valve is operated by a membrane. Fig. 11 shows the arrangement schematically. The diaphragm is exposed on one side to the Comprex air discharge pressure and on the other side to the exhaust manifold pressure. If at any instant the volume flow demand of the engine needs the available flow of the Comprex (which would appear by a pressure $P_{AM} = P_{EM}$, the balanced pressure on the diaphragm would close the throttle until $P_{AM} > P_{EM}$. The control system with bypass has been found suitable under all starting conditions so far experienced. If by the throttling, P_{AM} remains above ambient pressure and the volume requirements are satisfied, the checkvalve remains closed. At idling (22% maximum engine speed) about 50% of the flow comes through the checkvalve and 50% is furnished by the Comprex. At no load and higher engine speeds (50–100%) the Comprex furnishes up to 70% of the required flow. For higher loads at idling speed there is no operating point where the Comprex can deliver the required flow. The maximum bmep at this condition is about 90 psi. Only above 45% engine speed (1050 rpm) is it possible to open the throttle so that the checkvalve closes. With a standard fuel pump setting for constant torque, the engine receives an excess of fuel at lower than 45% engine speed. As a result, there is considerable smoke. As soon as this engine speed is reached, the Comprex timing is restored; the air manifold pressure increases instantly and full torque can be obtained.

In Reiners and Schwab[6] the similar problem with the turbochargers is described. At low speed and low load the turbocharger is idling and the engine can, at an instant call for power, reach naturally aspirated power. The engine and the turbocharger then have to accelerate before the manifold pressure can build up, since even at steady state below 40% engine speed it is not possible to obtain effective manifold pressures. In this case a fuel limiting system taking the manifold pressure as the signal accommodates this fact well. In the Comprex, where the limit is due only to mistiming below 40% engine speed, a fuel schedule depending on speed such as shown in Fig. 12 will eliminate smoke without any loss of response.

Starting on Grade — The inability to produce a high torque at low engine speed, both in the case of the Comprex and the turbosupercharged engine, represents a disadvantage in the case of starting on a grade. It is desirable to engage the clutch at 700 rpm (30% of maximum engine speed) to avoid excessive clutch wear. At this speed the engine produces a bmep of 90 psi, which is 35 hp at the engine or 30 road hp. The truck speed with fully engaged clutch in first gear is 1.9 mph with the gear ratios shown in Fig. 10. This power is sufficient to maintain constant speed at a 10% grade. If, however, the truck has to be accelerated from standstill, more power is required for acceleration. Assuming 2 sec for the acceleration, 10 extra hp are needed or 20 hp are available to overcome the grade. According to Fig. 10, this is a grade of 7%. If, however, full torque (bmep of 150 psi) is available, the road hp is 51 hp, which permits starting on a 12% grade. The further extension of the high torque to lower rpm might well be possible from the standpoint of air supply by the Comprex. It could well present a more difficult problem for the engine.

Noise — The high-frequency noise of the Comprex is considerably different from the standard diesel-engine noise. Each time a channel is opened at the exhaust port pulsations travel as sound waves through the exhaust pipe. The frequency established by rotor speed times blade number is a high sounding note. Mufflers of very simple construction which do not add any pressure drop have been found to attenuate the exhaust noise effectively.

In the intake, standard oil bath air cleaners have been found to be acceptable from the standpoint of intake noise attenuation and pressure drop. It is essential that metal pipes be used between the Comprex intake and the air cleaner to prevent trans-

Fig. 13 — Comprex supercharged truck used for road evaluation

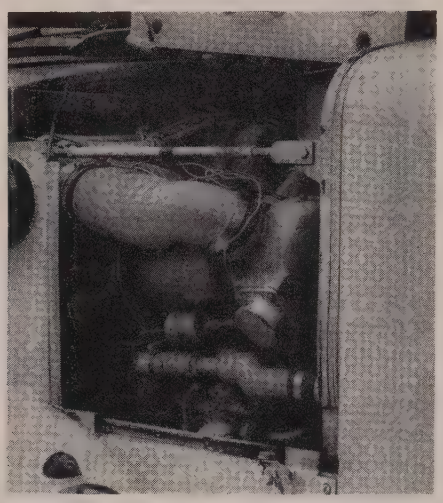


Fig. 14 — Comprex installation on truck

mission of noise directly to the outside. Any flat dome on the air cleaner through which sound energy can pass has to be lined internally with soft material. The noise of the powerplant in the experimental truck is very similar to the 2-stroke engine with Roots blower and can, therefore, be considered acceptable.

Truck Installation

A White tractor Model 9000 with a Cummins JT-6-B-20° diesel engine with Comprex supercharger used for the road evaluation is shown in Fig. 13 with the trailer used as test load.

The Comprex, being relatively small in diameter, can be mounted on an engine without the silhouette protruding excessively. The necessity of a drive imposes limitations to the possible location of mounting in many applications. A parallel arrangement to the engine centerline usually works out as the most satisfactory version, inasmuch as the Comprex can be driven from one of the available accessory drives on the engine. The Comprex is supported by the engine manifold which eliminates the need of hot pressurized expansion joints.

In the present truck installation (Fig. 14), dual air intake pipes are used for road evaluation. The new intake stator will need only a single pipe connection to the air cleaner. The air cleaner and connecting pipes to the engine are selected to limit the pressure drop to 6 in. of H_2O. The present exhaust system maintains a back pressure on the Comprex of about 8 in. of H_2O. It should be kept in mind that requirements for the low-pressure intake and exhaust are 20% larger than the engine flow, since some air blows directly through the Comprex. This has a disadvantage with respect to size. However, it helps to keep the rotor temperature as well as the muffler temperature low.

Intercooling — Considerable experience has been accumulated on dynamometer tests with intercooling. The matching procedures developed can be used as well for this case. The Comprex can easily produce high pressure; however, it has the disadvantage of higher air-compression and temperatures. Having a high air pressure and temperature makes intercooling very effective. Both air-to-air as well as air-to-water intercoolers have been tested and found to work very satisfactorily. It is easily possible to increase the engine bmep to 230 psi at the torque peak and about 200 psi at maximum speed.

Summary

The Comprex supercharger, using a new principle to utilize exhaust gas energy for the precompression of air, gives desirable operating characteristics for a truck diesel engine. The Comprex is a mechanically simple and rugged device tested over 20,000 miles. Response, clean exhaust, fewer gearshifts and improved fuel consumption are the predominant advantages offered.

DISCUSSION

Comprex Tested On Large Engine

— A. O. Kramer
Caterpillar Tractor Co.

AT CATERPILLAR, we have done some test work with the Comprex as a supercharger on a diesel engine. We used a Comprex approximately twice the capacity of the one discussed in the paper. My remarks will, of necessity pertain to this larger unit.

In a previous paper one of the authors presented the Comprex as being driven from the engine with a variable speed drive. In this paper the authors present the Comprex as having an improved speed range and being driven with a fixed ratio drive. I would like to add our observations on the relationship between engine speed and Comprex speed and some of its ramifications. The data shown on the following curves were collected by varying the Comprex speed while maintaining the engine load and speed constant. The Comprex used was one with the improved speed range. In Fig. A the boost pressure is maximum at a particular Comprex speed. It then falls off as the speed increases or decreases from this point. The differential between boost and back pressure is maximum at close to the same speed that gives maximum boost, and this differential decreases as the speed varies, although there is a fairly broad speed range with a substantial differential. The exhaust temperature is minimum at about the same speed that gives maximum boost, and rises with a

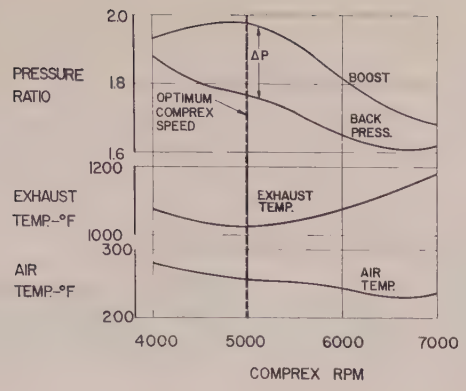


Fig. A — Effects of Comprex speed, constant engine load and

change of speed. The temperature of the compressed air does not decrease as would be expected with a decrease in boost, which indicates less efficient compression, or more heat transfer, or more mixing of the exhaust gas and fresh air, or a combination of the three. It is apparent that there is a best, or optimum, Comprex speed for the given engine load and speed. In this case the optimum Comprex speed is close to 5000 rpm.

It should also be mentioned that the allowable restriction on the intake and exhaust decreases considerably as the Comprex speed varies from this optimum speed.

These curves are quite representative of curves that would be obtained at any engine load and speed. The absolute values will change and particulars of the shape of the curves will change, but they will retain the characteristics shown. The optimum Comprex speed will shift slightly with engine load and speed, but for engine speeds from rated speed down to 40% of rated speed, and loads from full load down to where the Comprex is cut out of the system, this shift will be a maximum of 500 rpm. In other words, for a full range of engine loads and speeds the Comprex operates best at approximately 4500–5500 rpm.

Now what is the significance of this? It means that, with a fixed ratio drive, at high engine speeds the Comprex must run faster than its optimum speed and the performance characteristics would be similar to those represented by the right end of these curves; and at low engine speeds the Comprex must run slower than its optimum speed and the performance characteristics would be represented by the left end of these curves, or even extrapolations beyond the ends of these curves, depending upon the speed range of the engine. The use of a variable speed drive, with a speed ratio range that would match the speed range of the engine, would permit the Comprex to operate at or near its optimum speed. The engine could be rated at a higher load, and starting and acceleration characteristics would be enhanced. As the authors have mentioned, reliable variable speed drives, of a sufficient speed ratio range, that do not waste too much horsepower at the higher engine speeds, are relatively expensive; however, even with the improved speed range Comprex, the use of a simple, inexpensive fixed ratio drive *does* require a sacrifice of engine output potential.

Our findings support the authors' contention that the Comprex supercharged engine gives better lug, or torque, characteristics, and better response to load, than an engine with a free-running turbocharger; that is, a turbocharger with no controls. An improvement in gradeability, similar to what the authors cite for the highway truck, was noticed on rubber-tired earthmoving equipment. However, we should keep in mind that there are ways of improving turbocharged engine performance; for example, the use of an exhaust bypass controller improves the torque characteristics, and lower rotating inertia reduces the response time. The differences between turbocharged and Comprex supercharged performance must continually be evaluated.

The authors have stated that the Comprex is a mechanically simple and rugged device. To avoid the possibility that this might leave the impression that the Comprex is substantially trouble-free, I think it fair that I mention that we have found the Comprex has its share of the mechanical difficulties and failures that sometimes plague new devices.

We have found that the Comprex is much more sensitive to restriction at the intake and exhaust than is the turbocharger. This puts more severe limitations on silencing devices.

The authors mention that the Comprex produces a high sounding note that is effectively attenuated by mufflers of very simple construction that do not add any pressure drop. We have found that the Comprex generates a loud, high-frequency noise that requires more than the simple silencing technique that the authors have indicated, particularly in view of the limitation on exhaust and intake restriction. Both the intake and exhaust must be silenced.

The authors have expressed optimism about the Comprex being able to supply air for higher torques at lower engine speed, but they express a concern about engine problems developing. We would welcome, with enthusiasm, higher boost pressures and wider speed ranges, and urge the authors to continue their development along these lines.

Good Results Obtained With Turbine-Driven Supercharger

— Paul M. Uitti

Thompson Ramo Wooldridge Inc.

ALTHOUGH THE DATA shown appears quite good it seems proper at this time to compare some of the features of the Comprex with the turbine-driven supercharger.

In addition to the advantages of size, weight, and general noise level offered with the turbocharger, the higher pressure ratios readily obtainable are not to be overlooked.

In our laboratory we had occasion to test the JT engine using our 400 Series turbocharger and found that by running to pressure ratios of over 2/1 we could easily obtain outputs as high as 250 hp while still maintaining a downward trend in the specific fuel consumption curve. This represented a 100% increase over the naturally aspirated power.

Similarly, at the established 175 hp rating, with air/fuel ratios of over 30/1, in addition to the advantages of lower exhaust temperatures, better acceleration characteristics, and higher altitude reserve, an increase in torque of better than 30% was possible at 75% of rated engine speed.

There is no question that with the continued development and the ever-increasing trend toward supercharging of the diesel engine, the day is rapidly approaching when the naturally aspirated engine will become economically impractical.

ORAL DISCUSSION

Reported by J. H. Carpenter

Lycoming Division, Avco Mfg. Corp.

Q. What degree of pollution of fresh air by exhaust is obtained at full load?

A. 10% extreme, 5% middle range.

Q. Does Comprex compensate for altitude?

A. Altitude compensation not necessary. It is automatic up to approximately 12,000 ft.

Q. Is a single, or dual, manifold better on the Comprex?

A. Two manifolds are best.

Q. How does the Comprex affect starting?

A. No difference than with conventional compression.

Q. Has Comprex been tried on 2-stroke engine with and without Rootes blower?

A. It has been tried both ways. Comprex can operate over entire speed range without Rootes blower.

Q. What about seals in the Comprex?

A. Sealing accomplished by clearance only.

THE NUCLEAR turboelectric ion propulsion system operates as follows:

1. A nuclear reactor supplies thermal energy to a vapor boiler.
2. The vapor drives a turbine, which operates an electric generator.
3. The electric power is used to ionize and accelerate the propellant — a liquid alkali metal — which produces the thrust.

Electric propulsion systems can be used in a wide variety of space missions, including satellite sustainers, lunar cargo transport, and interplanetary missions.*

ELECTRIC

DURING THE NEXT decade we can expect the development of electric propulsion systems to result in the most practical, efficient method of propelling space vehicles over long distances in the solar system. In order to indicate the potential of these powerplants, it is interesting to examine the parameters affecting performance and component design of one type of system. In its simplest form these powerplants consist of an energy source, power conversion equipment, and electric rocket. Fig. 1 illustrates the many possible combinations of these three units which may be used together to form an electric propulsion system. Of the combinations listed, we shall examine the characteristics of the nuclear turboelectric ion propulsion system, which is one of the most promising ways to produce thrust for space propulsion.

Nuclear Turboelectric Propulsion System

One concept of an unmanned nuclear turboelectric propulsion system is illustrated in Fig. 2. A sodium-cooled zirconium hydride moderated reactor supplies thermal energy to a mercury vapor boiler. The wet vapor, about 30% quality, is passed through a separator to give saturated vapor at the turbine inlet. The vapor, after passing through the turbine, is condensed and subcooled directly in a tube-fin radiator. The vapor header in the center of the radiator, and the condenser lines serve as the main structural members. The aluminum fins are positioned in the lattice work of the condenser tubes and the headers so that they do not serve as structural members. The geometry of the radiator is dictated in large part by the allowable pressure drop, the need for extremely low weight construction and high reliability. For large powerplants, it may be desirable to condense the vapor in a compact indirect condenser unit where the waste heat from the condenser is transferred to a separate radiator by means of a liquid coolant circuit. The indirect condenser relaxes restrictions on radiator geometry, allowing a lower aspect ratio and hence requiring less ingenuity in radiator-vehicle integration.

Electric power is produced by a high-speed generator and is transformed into the required form with impedance matching equipment. The generator illustrated is an electromagnetic machine which has a large length to diameter ratio in order to keep weight to a minimum. For missions which dictate thrust chambers requiring high voltage, it is desirable to generate power with an electrostatic generator. It is attractive because of its higher efficiency and generally lower specific weight.

Nuclear radiation shielding is used to reduce reactor and boiler radiation to an acceptable level at the payload. In Fig. 2, lithium hydride is used for neutron shielding and depleted uranium for gamma shielding. These materials allow shield operation at high temperature where the energy absorbed in the shield is radiated directly into space. Total weight charged to shielding is kept to a minimum by using a lightweight separating structure between the power plant and payload for partial attenuation of radiation.

The electrical power supplied to the electric rocket is used to produce thrust by ionizing and accelerating the propellant. Fig. 3 illustrates the ion rocket operation. The propellant, a liquid alkali metal such as potassium, is vaporized and passed through a porous tungsten body at approximately 1600 F. The potassium vapor when in contact with the tungsten surface, under appropriate conditions, loses electrons and becomes ionized. The ionization process can be almost 100% efficient. After ionization, the potassium is accelerated to a velocity of the order of 100,000 fps by the electric field between the electrodes. The accelerator electrode is shaped so that the ions are focused through the grid. As the jet of ionized potassium is ejected beyond the second electrode, it is neutralized by an

* Paper presented at SAE National Aeronautic Meeting, New York, April 3, 1959.

PROPULSION SYSTEMS

Victor P. Kovacik and Daniel P. Ross
Thompson Ramo Wooldridge Inc.

electron emitting grid. The electrons prevent a continual buildup of positive charges on the powerplant.

Electric propulsions systems have unique characteristics when compared to chemical rockets. The comparison of the characteristics of electric rockets with chemical rockets is given in Table 1.

In the chemical rocket, energy is generated directly in the propellant by combustion. In the electric propulsion system energy is generated externally and then added to the propellant. The thrust of the electric rocket is generally a few pounds. The propellant flow may be a few pounds per day, however, with an exhaust velocity of 100,000 fps. The acceleration obtainable is about 0.001 that of gravity. Obviously, this powerplant cannot be used in accelerating masses through strong gravitation fields near the surface of planets. The necessary operating time for interplanetary missions can be several

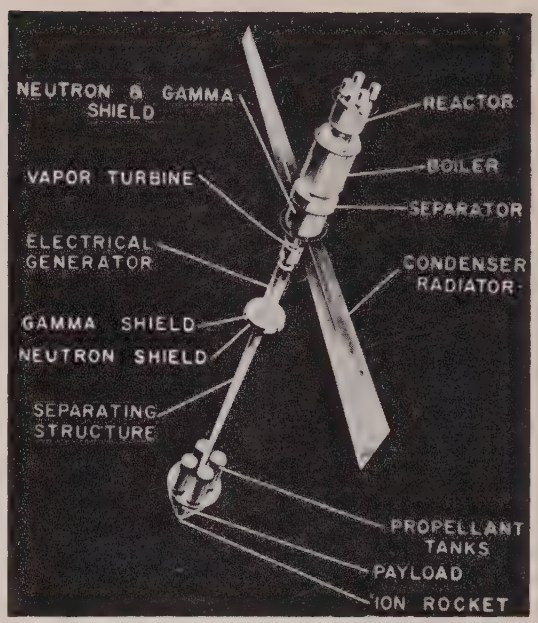


Fig. 2 — Nuclear turboelectric ion propulsion system

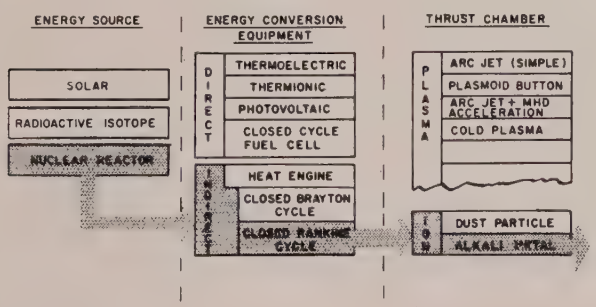


Fig. 1 — Segments of electric propulsion system

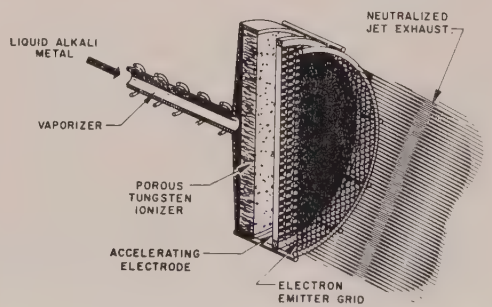


Fig. 3 — Alkali metal ion rocket

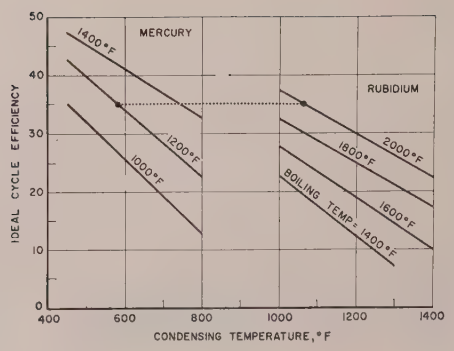


Fig. 4 — Efficiency of mercury and rubidium Rankine cycle

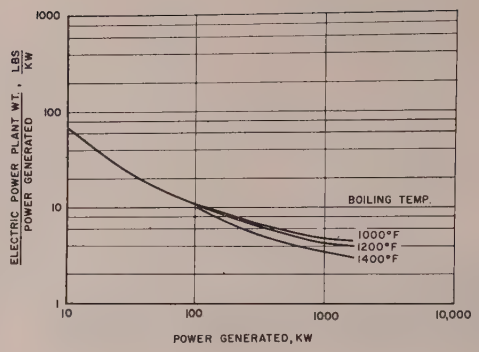


Fig. 6 — Specific weight of mercury vapor powerplant

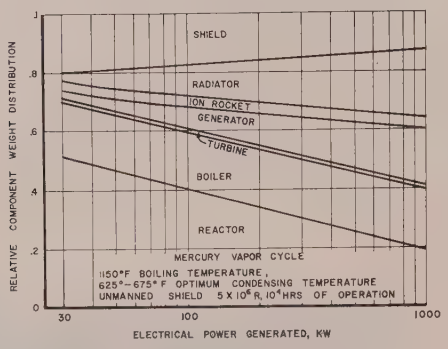


Fig. 5 — Distribution of component weight with electric power

Table 1 — Characteristics of Electric and Chemical Rockets

	Electric	Chemical
Propellant	Alkali metal	Hydrogen compound
Thrust, lb	0.01–10	1,000–1,000,000
Propellant Flow, lb/sec	0.00001	1,000
Exhaust Velocity, fps	10,000–1,000,000	10,000
Operating Life, hr	100–10,000	0.01–0.1
Rocket Weight / Thrust	1,000	0.001–0.01

thousand hours, and hence system reliability is extremely important.

Powerplant Cycle

It has been shown that the Rankine cycle powerplant is a particularly attractive method of generating power for space applications when considering the heat engine.[1] Mercury and rubidium are attractive as working fluids because of thermodynamic properties which match the requirements of a space environment. With each fluid it is possible to radiate waste heat at a relatively high temperature at a reasonable pressure. Also, it is practical to have a high maximum cycle temperature at a reasonable pressure, to give attractive cycle efficiencies. Practical boiling temperatures for the mercury engine are 1000–1400 F, while boiling temperatures for rubidium are 1700–2000 F. Optimum condensing temperatures for mercury engines are 600–700 F and for rubidium 1100–1200 F.

Fig. 4 gives typical values of ideal cycle efficiency versus boiling and condensing temperatures for both mercury and rubidium. Ideal efficiencies for the mercury Rankine cycle is near the upper limit of what can be attained with present state-of-the-art temperatures. The ideal cycle efficiency is approximately 94% of the Carnot efficiency. At the same efficiency the rubidium engine must operate at a maximum temperature about 600 F higher than the mercury engine, as indicated in Fig. 4. However, the higher boiling temperature permits a radiator temperature of approximately 600 F higher. This higher sink temperature enables the rubidium engine to operate with a small fraction of the radiator area, (but not necessarily weight) of the mercury engine.

The choice between mercury and rubidium depends upon the size of the powerplant, the required operating life, and what we think the state of the art will be in compact reactors, turbomachinery, and high-temperature structures. If one assumes that the additional 600 F maximum temperature required for the rubidium cycle can be met, the rubidium cycle has an advantage in geometry, particularly at powers above 100 kw where the radiator surface area is large.

Powerplant Weight

The operating power of the system dictates the relative importance of each powerplant component. If one considers the distribution of component weight given in Fig. 5, it is evident that the reactor weight and shield weight is of great importance at low powers. Below 20 kw of electrical power the reactor is greater than 60% of the total propulsion system weight. If we are to improve system weight at this power level, most gain is made by reducing reactor weight. As power level is increased, the boiler, generator, and radiator contribute a much more significant part of the system weight.

Fig. 6 gives the variation of specific weight of the mercury vapor system with power. At very large power, the powerplant weight is essentially that of radiator, boiler, and generator, and the powerplant specific weight approaches a constant value, if meteorite protection on the radiator is neglected.

At powers below 200 kw there is little incentive to go to temperatures much higher than 1000 F for the mercury vapor engine. The increased reactor, boiler, and turbine weight with higher temperature offsets the saving in radiator weight. At powers

[1] "Performance of Nuclear Electric Propulsion Systems," by V. P. Kovacik and D. P. Ross. Paper presented at IAS Annual Meeting, New York, January, 1959. IAS Report No. 59-25.

greater than 1000 kw, the specific weight of the powerplant can be reduced about 30% by going from 1000 F to 1400 F boiling temperature.

Mercury and Rubidium Vapor Power Cycles

The rubidium powerplant possesses no major weight advantages over the mercury powerplant. Fig. 7 gives the variation of specific weight with power for both the mercury and rubidium vapor systems. Even though the radiator area of the rubidium system is only a small fraction of that of the mercury system, the operating temperature of this radiator dictates designs which offer little weight advantage with respect to the radiator of the mercury system. On the other hand, from the standpoint of nuclear reactor technology, it appears that the boiling temperatures of the rubidium system require the use of basically heavier reactor design such as a reactor moderated by beryllium oxide. Even if the rubidium system utilizes a reactor as lightweight as the zirconium hydride moderated reactor, there is no significant weight advantage.

Fig. 8 gives a comparison of geometry for a 100-kw mercury vapor engine with a rubidium vapor engine of the same power. The advantage offered by the rubidium system is a favorable geometry by smaller radiator area. The mercury vapor engine offers favorable performance at mercury boiling temperatures almost consistent with those for existing long-life stationary mercury powerplants. Rubidium powerplants show some promise for applications requiring power outputs much greater than 200 kw. However, this requires that advances in high-temperature technology will permit the rubidium system to operate reliably for long periods of time, such as one year or more.

Space Missions of Electric Propulsion Systems

Electric propulsion systems can be used in a wide variety of space missions including satellite sustainers, lunar cargo transport, and interplanetary missions.[1] For example, electric propulsion systems appear attractive for a Martian interplanetary trip. For this mission, as well as other interplanetary missions, payload capabilities are maximized by using an electric rocket with a variable exhaust velocity.[2]

Fig. 9 gives the payload and power requirements of a 10,000-lb unmanned ion propelled vehicle making a round trip between the earth orbit and Mars orbit. A mercury vapor power system operating at a boiling temperature of 1150 F is utilized for the vehicle. It is assumed that the ion rocket efficiency of 70% remains constant with all variations in exhaust velocity which take place in the programmed flight. This vehicle, of which 5150 lb is payload, can travel from the earth's orbit to the Martian orbit and return in a period of 12 months of travel time. The electrical power output requirements for this payload are 275 kw.

A comparison of an electric propulsion system with a chemical system is given in Fig. 10 for transporting a payload of 5150 lb from the surface of the earth to a Martian orbit and back to the earth's surface with a travel time of one year. This neglects

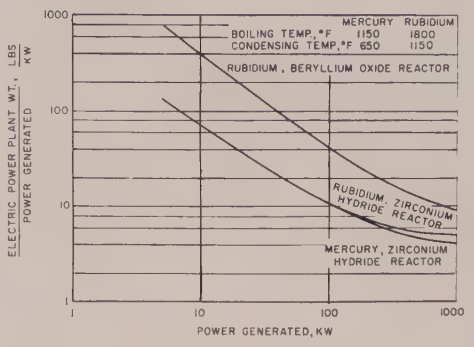


Fig. 7 — Comparison of specific powerplant weight

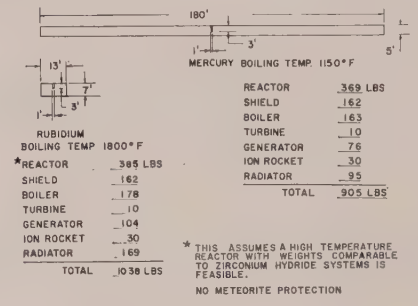


Fig. 8 — 100-kw mercury and rubidium propulsion system

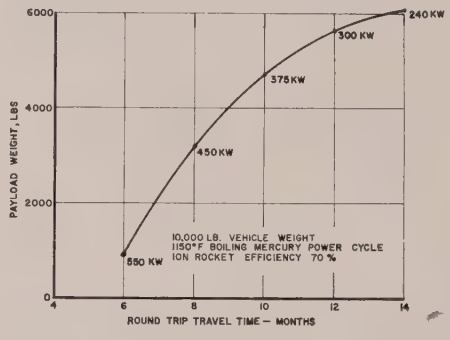


Fig. 9 — Unmanned ion propelled vehicle performance for earth orbit — Mars orbit — earth orbit mission

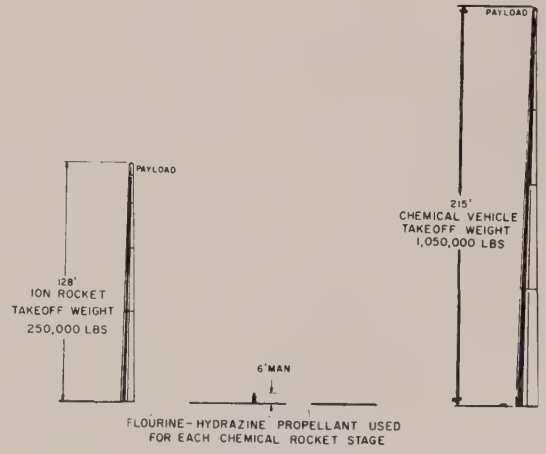


Fig. 10 — Comparison of chemical ballistic missile with ion rocket from earth's surface to Mars orbit and return (round trip travel time 12 months, payload of 5150 lb)

[2] "Comparative Performance of Ballistic and Low-Thrust Vehicles for Flight to Mars," by J. H. Irving and E. K. Blum. Paper presented at Second Annual AFOSR Astronautics Symposium, Denver, April, 1958.

the weight of the atmospheric braking equipment used in landing on the earth's surface. With conventionally fueled rockets, the electrically propelled vehicle requires a take-off weight of about 928,000 lb while an all chemical system requires about 8,000,000 lb, using weight estimates for chemical rockets given by Dobrin.[3] The chemical boosters utilize oxygen-jet fuel, Nitric Acid-UDMH, and high-performance solid propellants. For high-impulse propellants, it is tacitly assumed that by the time large power electric propulsion systems are operational, a rocket having the performance of a fluorine-hydrazine rocket will be available. The take-off weight in this case is 250,000 lb for the chemical-electric rocket, and 1,050,000 lb for the all chemical rocket.

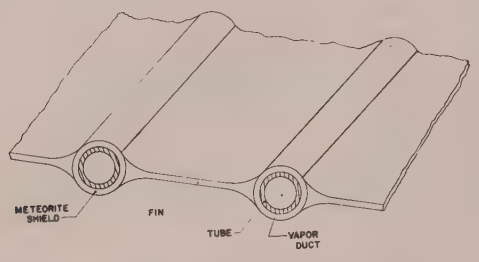


Fig. A — Fin tube geometry

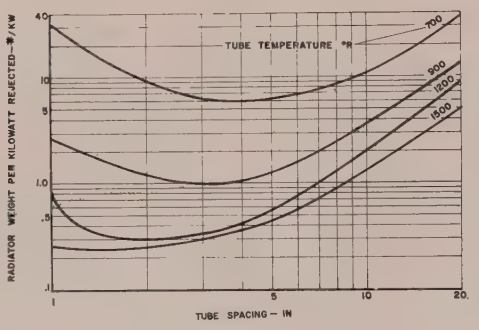


Fig. B — Radiator weight per kilowatt rejected (meteorite shielding included)

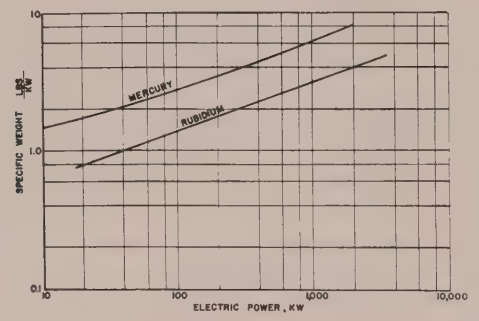


Fig. C — Radiator specific weight versus power level (50% probability of penetration in two years)

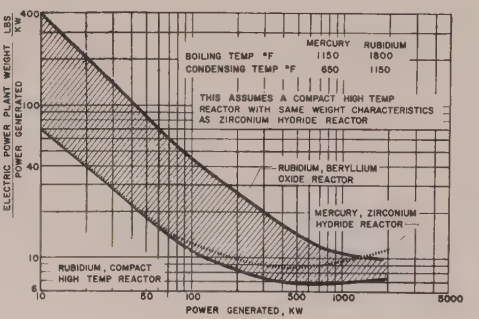

ELECTRIC POWER PLANT WEIGHT LBS/KW POWER GENERATED
MERCURY
RUBIDIUM
BOILING TEMP °F 1150 1800
CONDENSING TEMP °F 650 1150
THIS ASSUMES A COMPACT HIGH TEMP REACTOR WITH SAME WEIGHT CHARACTERISTICS AS ZIRCONIUM HYDRIDE REACTOR
RUBIDIUM, BERYLLIUM OXIDE REACTOR
MERCURY, ZIRCONIUM HYDRIDE REACTOR
RUBIDIUM, COMPACT HIGH TEMP REACTOR
POWER GENERATED, KW

Fig. D — Comparison of specific powerplant weight (with meteorite shielding, 50% probability of penetration in two years)

Discussion[4]

The performance data presented in this paper utilizes lightweight radiator concepts developed at Thompson Ramo Wooldridge, Inc. The tube-fin design illustrated in Fig. A gives very favorable performance for mercury vapor systems. In this concept, foil thick aluminum fins, which serve primarily as radiating surface, carry the heat away from the tubes. The meteorite shielding is only required around the tube in which the vapor condenses and not on the radiating fins. For minimum radiator weight, the temperature distribution within the fin must be linear. this requires fin shapes which are approximately triangular.[1] Furthermore, for a given tube size there exists an optimum tube spacing for a given tube wall temperature. (See Fig. B.[5]) In general, a fin effectiveness between 0.3 and 0.5 results in minimum radiator weight. However, total radiator area is two to three times that of an isothermal plate-type radiator operating at the tube wall temperature. Computations indicate that a decrease in tube diameter results in a large decrease in radiator weight. Therefore, tube diameters are selected as small as possible, limited by pressure drop and other practical considerations.

The radiator designs given above (Fig. A) include meteorite shielding weight. However, the weight data presented in this paper does not include that of meteorite shielding. The effect of meteorite shielding on the tube-fin radiator is shown in Fig. C.[5] A 50% probability of meteorite penetration over a 2-year period is used for the criteria for the specific weight of the radiator. The specific weight of the radiator shielding against meteorites increases with power due to the larger target area to the meteorites. For the rubidium vapor system described in the paper, an isothermal steel plate condenser-radiator was considered as the operating temperatures are too high for designs employing aluminum fins. If meteorite shielding is to be used with the plate design, a poor specific weight results as the meteorite protection must cover the entire plate. Our investigations of the high-temperature steel tube-beryllium fin radiator indicate more favorable performance than the plate radiator, for the rubidium powerplant. Fig. C gives the specific weight of this steel tube-beryllium fin radiator, including the weight of meteorite shielding.

Specific weights of mercury and rubidium systems which include the effect of meteorite shielding on tube-and-fin radiator concepts are given in Fig. D. For both the mercury and rubidium powerplants there is a minimum point in the specific weight because meteorite protection is included. The weight penalty for meteorite protection indicated in Fig. D shows the importance of the radiator for powers greater than a few hundred killowatts. Without meteorite protection the specific weight of each powerplant continually decreases with power, becoming asymptotic to a constant value. (See Figs. 6 and 7.) The two curves for the rubidium system in Fig. D place the bounds of its specific weight depending on reactor developments. Most probably a compact reactor development for the rubidium system will involve some high-temperature beryllium or hydrogen compound which will require larger reactor weights than the zirconium hydride system but lower than the beryllium oxide system. For this case the rubidium power system will become lighter than the mercury power system at a power between 500 and 1000 kw. Nevertheless, for power requirements less than 1000 kw, the rubidium power system offers little weight advantage over the mercury power system, whether or not appreciable meteorite shielding is included.

[3] "Estimating Missile Proportions," by Dr. Saxe Dobrin. *Aircraft and Missiles Manufacturing*, Vol. 1, April, 1958.
[4] To clarify points on radiator design raised at time of paper's presentation.
[5] Based on the research of R. G. Gido of Thompson Ramo Wooldridge Inc.

Engine Blocks and their Components in *ALUMINUM DIE CASTING*

A. F. Bauer
Doehler-Jarvis Division, National Lead Co.

THE LIGHT WEIGHT of aluminum has always intrigued the automotive engineer and he has considered its use ever since automobiles have been built. However, these early efforts failed because of the high initial metal costs of aluminum. Since aluminum sand casting — as the only casting method then available for large parts — could not offer any foundry advantages over gray iron which would have compensated for its high metal costs, engine blocks and other large engine castings remained in gray iron. For a long time gray iron maintained its strong position as the major casting material for engines, not because of its superior properties but because of its lower manufacturing costs.

This picture changed when aluminum's biggest advantage over gray iron, its adaptability to modern fast-cycling casting methods, such as permanent mold and die casting, was fully recognized and materialized. In the die casting process the higher metal price of aluminum is more than compensated by the conversion of molten metal in one single operation into a highly finished casting of high properties, smooth surface, and high accuracy. Because of their fast production rates and reduced machining costs, aluminum die castings became competitive with gray iron in many cases more than 20 years ago. But in these early stages the size and quality of aluminum die castings were limited by the lack of know-how and the size of the existing die casting machines.

As the die casting technique improved and larger die casting machines were developed, aluminum die castings continued their successful penetration of the automotive industry. Intricate valve bodies for automatic transmissions were followed by highly stressed large parts, such as clutch housings, torque converter housings, and transmission cases. This development was climaxed by the production of the first die cast 6-cyl engine block in 1955.

While die cast aluminum engine blocks are still not used on a mass-production basis, it must be said that the pioneer work of the die casting industry and the availability of large die casting machines has had a distinct effect on the inroad aluminum has made in the automotive industry during the last 5 years.

Large Die Casting Machines

The story of the first die cast engine block for passenger cars begins in April, 1951, when a joint

WITH IMPROVED production techniques, aluminum die castings have become more and more attractive to automotive engineers — to the point that the author expects mass production of die cast water-cooled engine blocks soon.

Conversion into aluminum die castings presents three major problems: (1) elimination of undercuts, (2) solution for a wear-resistant cylinder bore, and (3) solution for the expansion problem of crankshaft bearings. This paper discusses several possible solutions.*

program for a 4-cyl engine block and the necessary die casting facilities was initiated in a meeting between the Kaiser Aluminum & Chemical Sales, Inc. and ther automotive plant Kaiser-Frazer on the one side, and the Doehler-Jarvis Division of National Lead Co. on the other side.

Since existing die casting facilities were much too small for this engine block, the design of the world's largest die casting machine was started in 1951. This machine (Fig. 1) was completed and made its first test run in October, 1954. It has a locking pressure of 2000 tons which was at that time more than twice the locking pressure of the largest die casting machine on the market. The tie bar distance is 72 in. × 72 in. and the tie bars are 1 ft in diameter. The metal is injected into the die with pressures between 4000 and 8000 psi. With a capacity for castings up to 75 lb. aluminum and platens large enough to accommodate the largest V-8 cylinder engine block for passenger cars now in use, this machine has retained its reputation as the world's largest die casting machine.

In the meantime, a number of large die casting machines has been designed and erected, closing the gap between the existing 800-ton presses and the Doehler-Jarvis 2000-ton machine. Machines with locking pressures of 1200 tons were set up to produce transmission housings up to 25 lb. Larger machines with 1400- and 1600-ton locking pressure capable of producing engine blocks up to 50 lb are under construction. The first of these *1600-ton* die casting machines — designed and built by Cast-Master, Inc. in Cleveland, Ohio, for the explicit purpose of producing a medium-size 6-cyl engine block — may soon make its first trial run at Chrysler's die casting plant in Kokomo.[1]

Today a great number of large die casting machines are available to take care of coming requirements in die cast engine blocks. The fast growth in the size of the machines is the best indication for these tremendous developments in the aluminum die casting industry.

First Die Cast Engine Block

As mentioned before, our work on the die cast engine block actually started in 1951, with a 4-cyl engine block of about 28 lb for Kaiser-Frazer. Several prototypes in aluminum sand casting identical to the die cast design were built and tested by Kaiser. The aluminum engine block stood up quite well and all novel design features introduced with aluminum die casting worked out very well on the sand cast prototype shown in Fig 2. The efficiency of the engine was beyond expectations and a cost comparison of the completely machined 4-cyl engine block revealed that the aluminum die cast engine block was appreciably cheaper than aluminum permanent mold and presented a saving over gray iron of more than 10%.

Before the die design for the 4-cyl engine block was completed, two more cylinders were added resulting in a 6-cyl engine block for a 200-hp engine. It was intended to produce these engines at a rate of 400 per day. But before the die was completed, basic changes in the Kaiser passenger-car program led to the cancellation of the whole project. Since the die casting die, however, was far advanced and about 75% completed, construction of the die was continued and the first die cast 6-cyl engine block produced in March, 1955 (Fig. 3).

This first die cast 6-cyl engine block weighs only 43 lb and is about 130 lb lighter than in gray iron. It contains 129 holes all of which are cored. Drilling operations, therefore, have been completely eliminated. Machine stock varies between 0.030 and 0.060 in. The wall thickness is uniform and with 5/32 appreciably thinner than in gray iron. This engine block, made in 1955, is still the largest aluminum die casting ever produced. But it will hold this record for only a few more months when heavier engine blocks will be produced.

None of these die cast engine blocks was completely machined or assembled. Our tests, therefore, were restricted to checking the quality of the part as a die casting. Actual field tests with engines using the die cast design with wet sleeves were confined to the 4-cyl aluminum sand cast engine (Fig. 2).

For us, who spent much time in the development of the first 6-cyl engine block, it was most regrettable that none of these die cast blocks ever was completely machined, assembled, and tested under field conditions. Such tests would have given us important information on the west sleeve sealing problem, the corrosion problem, the expansion problem, the strength and rigidity of an aluminum block, the noise problem, and other objections which we have to answer daily.

What the die casting of the first engine block accomplished is to prove that such large intricate castings can be successfully produced in aluminum die casting, to confirm the many advantages they

* Paper presented at a meeting of the South Bay Division, SAE Northern California Section, San Francisco, Feb. 3, 1959.

[1] Pilot runs of die cast 6-cyl engine blocks have in the meantime been successfully produced at the Kokomo plant of Chrysler and at Doehler-Jarvis, Toledo.

have over other casting methods, and to encourage others to go to die casting designs on future engine blocks. And judging by the present activities, I believe that this has been accomplished.

The basic changes in the redesign of the engine blocks to die casting may best be explained in the isometric drawing (Fig. 4). The first basic change from gray iron to aluminum is the wet sleeve arrangement necessary to form the pockets for the water-cooling system. The six wet sleeves are placed into the bores of the aluminum cylinder block and sealed against the cooling water. This method of wet sleeves was not well accepted by the automotive industry, which for many reasons gives preference to dry sleeve arrangements.

The second basic change was the relocation of the camshaft from the conventional spot on the side of the engine block to the top of the cylinder head. This relocation was done to simplify the engine block as a die casting. However, we feel that this is no longer necessary, and our latest die casting designs for engine blocks leave the camshaft in its standard location.

The third change refers to the deep pockets on both sides of the block. These pockets are actually metal savers to eliminate heavy metal sections behind the main bearings as shown in Fig. 5. This picture also shows the long oil holes for the main bearings, connecting them with the oil channel which is cast as an open channel on the side and closed and sealed by a cover. Today, only four years later, we consider this design obsolete in

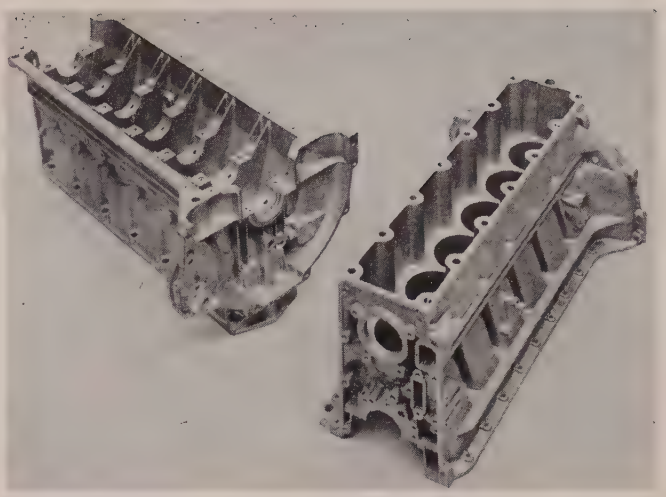


Fig. 3 — Six-cyl die cast engine block

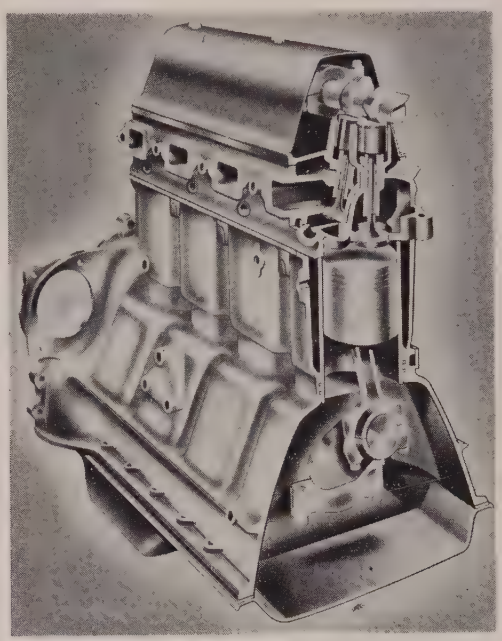


Fig. 4 — Isometric drawing of cylinder block

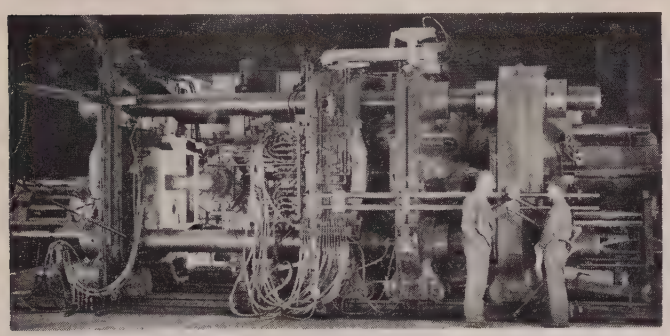


Fig. 1 — 2000-ton die-casting machine

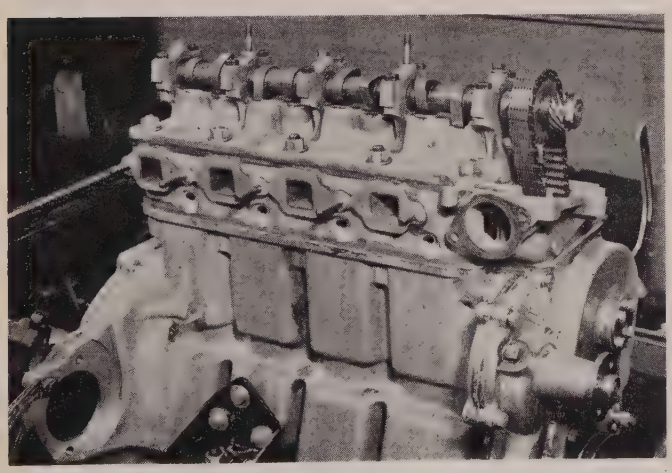


Fig. 2 — Four-cyl engine (aluminum sand casting)

Fig. 5 — Cross-section of 6-cyl block

many respects. The latest design of engine blocks necessitates fewer basic changes.

Major Conversion Problems

When we had die cast the first 6-cyl engine block in 1955, nobody was then interested in 6-cyl engines. The V-8 engine was the only thing three years ago. We, therefore, designed V-8 cyl engine blocks with wet sleeve arrangements. One of the many solutions we had worked out is shown in Fig. 6. In this engine block the undercut problem is solved by designing the camshaft and valve tappet bearing into a separate die casting which, fastened to the block as shown, builds the oil channels. While many companies studied the details of our V-8 cyl design and accepted its basic features, the cost analysis revealed in several cases that the wet sleeve arrangement is very expensive and offsets all the cost advantages the die cast aluminum block can offer. Furthermore, fear of water leakage has affected the adoption of the wet sleeve design. In addition, electrolytic corrosion between the two dissimilar metals was also considered to be a serious problem; and last, the wet sleeve arrangement presents an assembly problem which would interrupt the smooth flow of the block through the transfer machining line.

Therefore, we started to work on cast-in sleeves, two versions of which are shown in Fig. 7. On the left side is shown a cast-in gray iron sleeve which is essentially still a wet sleeve. Only a small portion of the sleeve is cast into the aluminum; therefore, this design is not too rigid. In addition, the two dissimilar metals forming the water channels present a problem of electrolytic corrosion.

The cylinder on the right side shows a cast-in gray iron sleeve with aluminum cast around it over its full length. Since water does not come in direct contact with the sleeve, this design is called the dry sleeve arrangement. This dry sleeve arrangement is today considered the safest and most economical solution.

After two years of concentrating on a die cast V-8 cyl design, the 6-cyl engine came to life again as the American automotive industry shifted their interest to smaller cars.

For us diecasters, the basic conversion problems from gray iron to aluminum die casting remain the same — whether it is a 4-, 6- or 8-cyl engine of the "V" type or straight in-line type, whether it's water-cooled or an aircooled type engine. This conversion into aluminum die castings presents three major problems which I would like to explain on Outboard Marine's V-4 engine block which is already in large production as an aluminum die casting. (See Fig 8.)

These three problems are:

1. Elimination of undercuts.
2. Solution for a wear-resistant cylinder bore.
3. Solution for the expansion problem of crankshaft bearings.

These are difficult problems for which solutions can be found only through close cooperation between engine designer and die casting expert.

Elimination of Undercuts — The elimination of undercuts and back drafts is, no doubt, the most serious of these three conversion problems. Since the metal is injected under extremely high pressures of up to 10,000 psi, the die casting has to be made out of heat-treated steel to withstand the impact of the injected metal. The cores must retract and the casting must be ejected from the die. Sand cores cannot be used and other destructible materials which would withstand the impact of the injected metal have not been found yet. Loose pieces to build undercuts are possible but usually retard the casting cycle and, therefore, increase the costs. In the picture of the V-4 engine block (Fig. 8) you see how the undercut problem of the water pockets has been solved by opening them to one side so that the cores can be withdrawn. The undercut in the center of the "V" section has been eliminated by big side cores. Slides in every direction resulted in uniform walls and, all together, in a most efficient design of a lightweight, high-strength engine block in aluminum for a 2-stroke outboard motor of 50 hp.

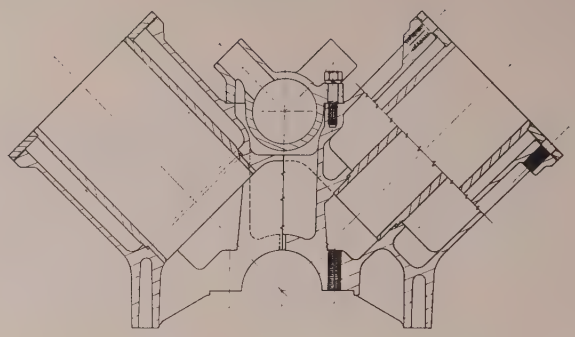


Fig 6 — V-8 cyl with wet sleeves

Of course, with the larger 6- and 8-cyl engines for passenger cars, the elimination of undercuts becomes a still more serious problem.

Many automotive engineers consider the necessity to design completely without undercuts as die casting's biggest disadvantage. The fact that die casting cannot resort to sand cores restricts, of course, the flexibility of design. But let us not forget that sand cores which have to be baked and then put carefully in the mold by hand, are extremely costly and increase casting costs, whether it is in sand casting or in semi-permanent mold casting.

And by the same token: castings without undercuts are less expensive to produce in any casting method and quite naturally lend themselves best to the fastest casting method, the die casting process.

In other words, the more we strain our brains in the search for a design without undercuts, the greater will be our reward in reduced manufacturing costs of such a part.

Dry Sleeve Arrangement — The 2-cyl engine block for a 40 hp outboard motor (Fig. 9) is another good example for the elimination of undercuts. However, it is shown mainly to demonstrate the ease of casting gray iron sleeves into aluminum die castings. Since unprotected aluminum does not have the wear resistance necessary to withstand all wear conditions, the cylinder bores of an aluminum engine must be protected in some way. One of the methods used is to equip the bore with a gray iron

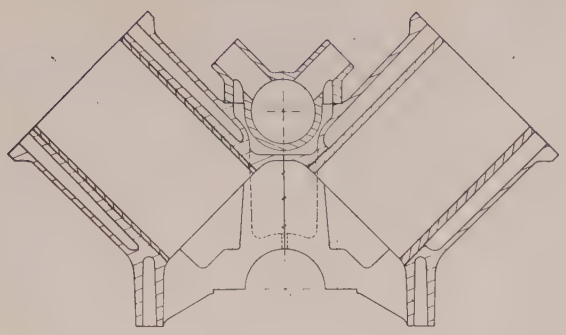


Fig. 7 — V-8 cyl engine block with cast-in gray iron sleeves

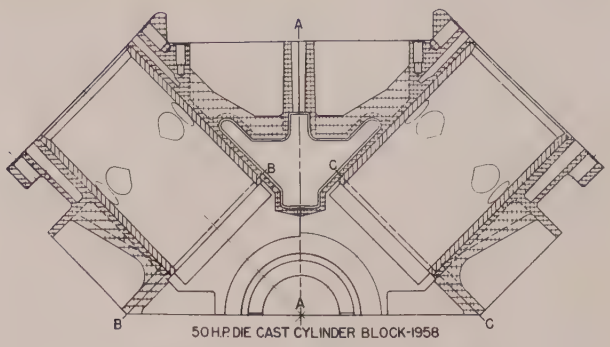


Fig. 8 — V-4 cyl block (cross-section)

liner. These liners can either be pressed in or cast in.

Cast-in gray iron sleeves are held only by the residual stresses which develop when the molten aluminum, injected under high pressure, cools off and shrinks onto the sleeve resulting in a strong mechanical bond with good thermal conductivity. These residual stresses have not been found detrimental to the design, and many thousands of outboard motors have used and are still using this arrangement quite successfully. However, sometimes the aluminum may shrink away from the liner on heavy sections. The voids, thus created, retard the flow and may cause "hot spots" in the sleeve, especially when they are located near the top. Such a void is shown in Fig. 10 on the left-hand side.

Since these shrinks have led to a few complaints from our customers, we have developed a new and improved method of bonding these cast-in sleeves to the aluminum. In this method, shown on the right side of Fig. 10, the die cast aluminum penetrates, under pressure, into the fine undercut-like crevices of the rough surface of a centrifugally cast sleeve, thus creating a bi-metallic interlock bond which is mechanical in nature but extremely stong.

The enlarged view, on the lower part of Fig. 10, shows how the interlock between the die cast aluminum and the spine-like particles of the gray iron is accomplished. It may be appropriate to insert here that all centrifugally made gray iron castings have this rough surface *on the outside diameter*. This roughness has been well-known and considered a serious drawback so much so that it has been generally removed by a tumbling action. To take advantage of the rough, as-cast surface of a centrifugal casting and to create an interlocking bond with the die cast aluminum, however, has been the development of our Research Department. Since the bond is a mechanical interlock of the two metals, we have given it the name B. M. I. bond which stands for *Bi-Me*tallic *I*nterlock bond. This method of B.M.I. bonded sleeves is not applicable to aluminum sand casting nor aluminum permanent mold casting, since the extremely small static pressure of the molten metal is not strong enough to force the aluminum into the fine crevices of the surface of the sleeve.

Tests with B.M.I. bonded gray iron sleeves in aluminum die castings have shown outstanding results. The heat conduction is better and the

Fig. 9 — Two-cyl engine block (outboard) with dry sleeves

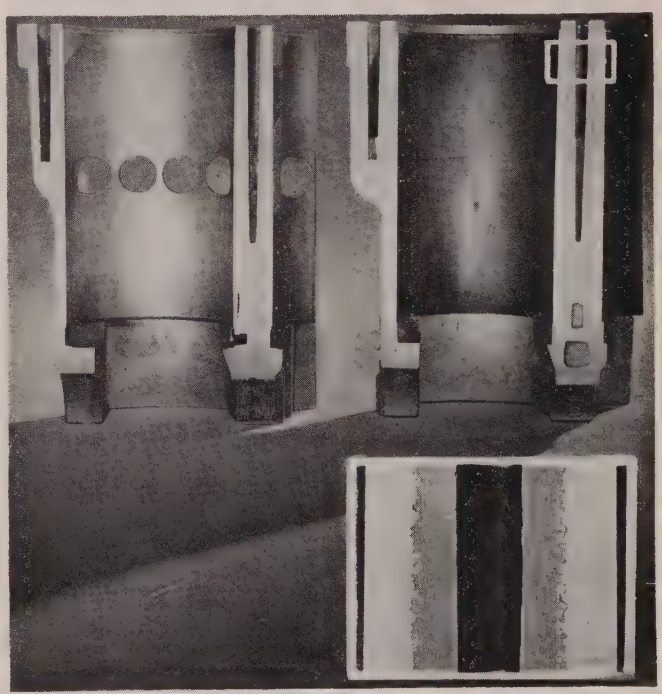


Fig. 10 — Outboard engine blocks — B.M.I. bonded cast-in sleeves

mechanical properties of the bond are higher than with any other bonding method between aluminum and iron. B.M.I. bonded samples which we have tested show a shear strength of up to 18,000 psi while plain cast-in sleeves stood only a shear load of 6000 psi. The B.M.I. bond creates such an intimate bond between the two metals that they do not separate at any temperature even when heated close to the melting point; while on plain cast-in samples after repeated heating cycles a separation between both metals may take place. The Doehler-Jarvis Research Lab has completed extensive tests with this bond and has published a report on its findings.[2]

The cast-in, B.M.I. bonded gray iron sleeve is, in our opinion, the most logical design and presents the best solution for die cast engine blocks because it is a method already proved on smaller engines. For this reason, the method of B.M.I bonded dry liners has been chosen for most of the die cast engine blocks in passenger cars now under consideration.

Wet Sleeves — In Fig. 11 is shown our first die cast engine block and the wet sleeve arrangement we had chosen because this design gave us the simplest die casting six years ago. While this simplification of the die casting was accomplished, the six gray iron sleeves, soldered into a steel deck plate, resulted in a very expensive design. As I have already mentioned, we believe today that the automotive industry does not go for wet, gray iron sleeves not only because of its high costs but also because of the galvanic corrosion problem of the two dissimilar metals. The interrupted machining and assembly of sleeves and blocks is also considered most undesirable.

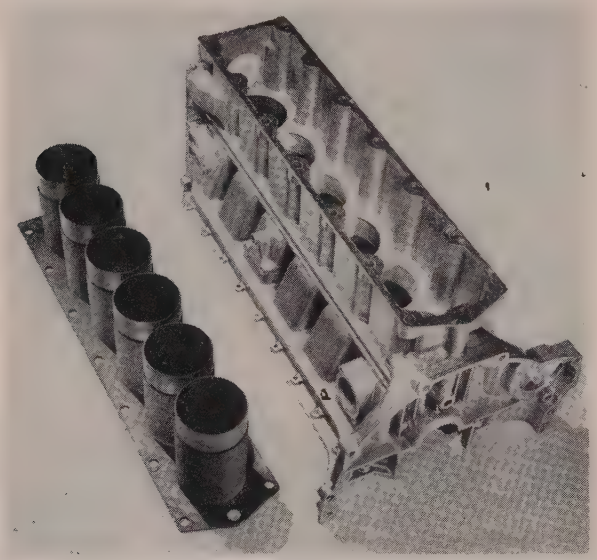


Fig. 11 — Six-cyl block with wet liners mounted in steel deck plate

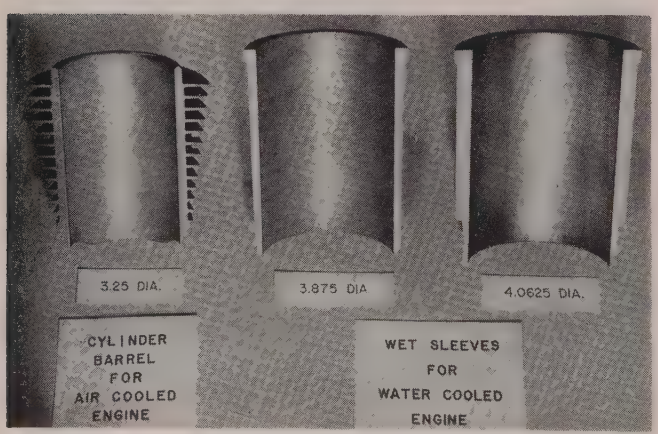


Fig. 12 — Transplant coated cylinder sleeves and barrels

Cast-iron sleeves in a steel deck plate, as shown in Fig. 11, result also in an additional expansion problem. A better solution is to use wet sleeves in aluminum. They have the advantages of lighter weight, better thermal conductivity, and the same coefficient of expansion as the piston, resulting in better engine efficiency and reduced oil consumption. But the bore of aluminum sleeves must be protected against excessive wear. Porous chrome plate and sprayed molybdenum-iron coats on aluminum have both shown satisfactory wear resistance. Chrome-plated aluminum sleeves are already successfully used in European cars. However, both methods of chrome plating or metal spraying of cylinder bores are considered to be much too expensive.

Transplant-Coated Cylinder Sleeves and Barrels — In the endeavor to find an economical solution to the sleeve problem, the Doehler-Jarvis Division has developed a new die casting method with which liners for water-cooled engines and cylinder barrels for aircooled engines can be produced with a wear-resistant metallic coat of extreme smoothness and high accuracy. Fig. 12 shows a cylinder barrel of an aircooled engine and two wet sleeves coated with 0.020-in. iron-coat by this new method. The coat is so thin that you may have difficulty detecting it in the picture.

This coat can be produced in almost any material, in any thickness and in any degree of porosity; therefore, it is most versatile. For economical reasons it is held relatively thin. Die cast sleeves and cylinder barrels with this transplant coat can be produced very economically on semi-automatic or automatic equipment and create a product of highest quality. The enlarged view (Fig. 13) shows that an excellent mechanical bond exists between the wear-resistant coat and the aluminum. We call this new coating method a transplant coat because it is first created on the core of the die and then transplanted to the die casting when the injected metal shrinks onto the rough outside of the sprayed coat. Taking advantage of differential expansion, the steel coat stays with the aluminum die casting and creates a taper-free bore of an extremely high smoothness of approximately 30 rms and a high accuracy of ± 0.003 in. The transplant coat (Fig. 12) is so smooth that slight honing is the only operation necessary to finish the bore.

Compared with the plating and standard spray coats on aluminum — where the cylinder bore first has to be machined to close tolerances, then coated and honed — this new transplant coat represents a tremendous savings because the coated bore of

[2] Research Bulletin No. 2, Doehler-Jarvis Division, National Lead Co.

high accuracy is actually produced in the same cycle as the die casting is created.

Since the coat is hard and porous, there are indications that it will retain oil for cold start conditions as good as gray iron does. Several test runs of transplant-coated sleeves are at present being conducted by a number of automotive companies, and first results indicate that the coat shows little wear and most satisfactory operating conditions.

If these tests continue to be as favorable as they are at the start, this transplant coat may present a better and more economical solution for cylinder barrels of aircooled engines than the present chrome-plated aluminum barrels. Even for those economy cars which now use the least expensive gray iron cylinder barrels — such as the Volkswagen and the new Chevrolet — this now design may present a welcome improvement in quality and manufacturing cost.

Transplant-coated aluminum die cast cylinder barrels for aircooled engines will definitely give better operating conditions at all temperature levels and, therefore, result in higher engine efficiency than gray iron cylinder barrels.

The chances for transplant-coated aluminum sleeves in water-cooled aluminum engines are extremely difficult to predict. While we are sure that die cast and transplant-coated aluminum sleeves, as shown here, are less expensive than gray iron sleeves and eliminate the electrolytic corrosion, the problem of assembly and sealing of each sleeve still remains and may be the main stumbling block against this design.

Expansion Problem on Crankshaft Bearings — Probably the most confusing of all the conversion problems is that of the differential expansion between aluminum and gray iron in the cylinder block. (Fig. 14) shows the coefficient of expansion of several aluminum die casting alloys in comparison to iron. Values of several hypereutectic Al-Si alloys have been added because the expansion decreases appreciably with increased Si content. The temperature range up to 300 F is the most critical one for our consideration.

At every point in the engine block where aluminum is in a dimensionally close relationship to iron parts, there is a problem of differential expansion. In the cast-in gray iron cylinder sleeves the differential expansion sets up residual stresses which, as experience has shown, are not detrimental to the design. The differential expansion over the length of the aluminum block can be controlled by allowing enough side clearance for crankshaft and connecting rods. For the relatively small camshaft bearings, the differential expansion does not create a problem. This leaves the bores of the crankshaft bearings as the only serious expansion problem.

With a euteltic Al-Si alloy (approximately 12% Si) which we prefer for engine blocks as the best compromise between castability, mechanical and physical properties, the bore of a crankshaft bearing of 2½ in. diameter theoretically expands 0.004 in. more than the crankshaft in temperature range of 300 F.

With so much additional clearance the engine becomes noticeably noisier. In our search for acceptable solutions, we investigated a number of European aluminum engines but found that they had done nothing to overcome this expansion problem in crankshaft hearings. Some companies in Europe resort to a larger oil pump and create a heavier oil film in the bearing which reduces the noise when the engine temperature increases and the bearings expand. In one of the latest American designs, large steel blocks were fastened into the aluminum engine block so that the bearing would not be affected by the higher coefficient of expansion of aluminum. This design overcomes the expansion problem all right, but is very expensive and much too heavy. Others just use gray iron caps and seem to get away with it.

A completely new approach is the Doehler-Jarvis design of an "Expansion Compensating Bearing Cap" shown in Fig. 15. This design works on the bimetal principle and compensates the bearing expansion by the residual stresses set up between the aluminum cap and the cast-on steel strip.

Minor Conversion Problems

Besides these three major conversion problems,

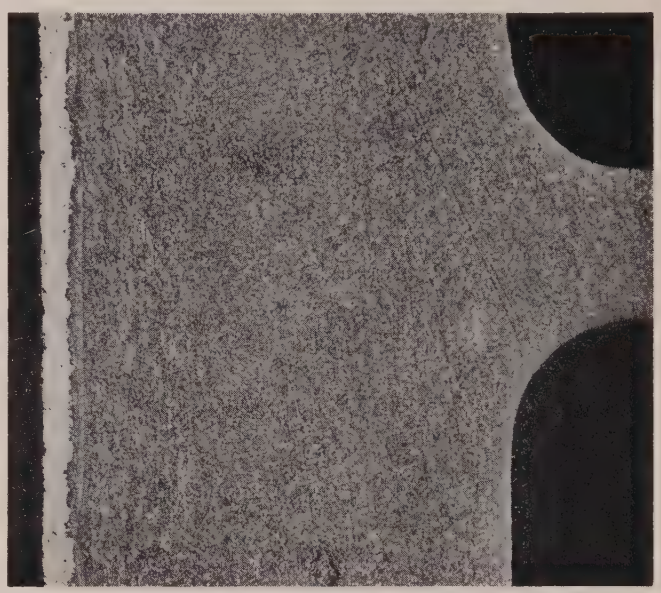


Fig. 13 — Transplant coated cylinder sleeves and barrels (enlargement)

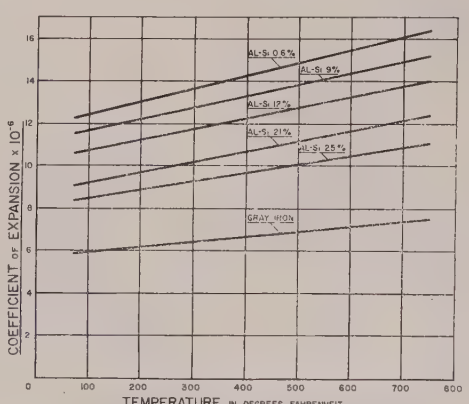


Fig. 14 — Coefficients of expansion of aluminum alloys versus gray iron

we experience a number of minor problems and objections which have to be overcome in the introduction of aluminum for engines.

Structural Strength of Aluminum and its Wall Thickness in Relation to Gray Iron — Some engineers have the mistaken idea that aluminum die castings are much weaker than gray iron castings and that their walls have to be increased so much that the weight advantage is somewhat reduced. We find by experience that the opposite is the case. Table 1 shows the weight comparison and the weight savings of a few engine blocks which we redesigned for aluminum.

Table 1 — Weight Comparison of Cast Iron versus Die Cast Aluminum

Type of Engine Block	Weight of Cast Iron Block as Cast, lb	Weight of Die Cast Aluminum Block, lb	Weight Savings, lb	Weight Ratio
6 cyl	173[a]	43	130	4/1
4 cyl	122[a]	28[a]	94	4.3/1
V-8 cyl	208	55[a]	153	3.8/1

[a] Calculated weights.

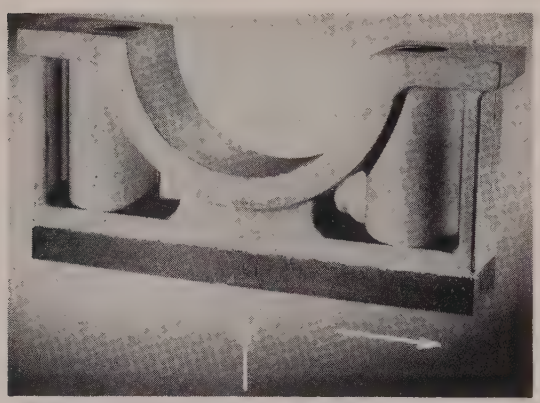


Fig. 15 — Bearing cap

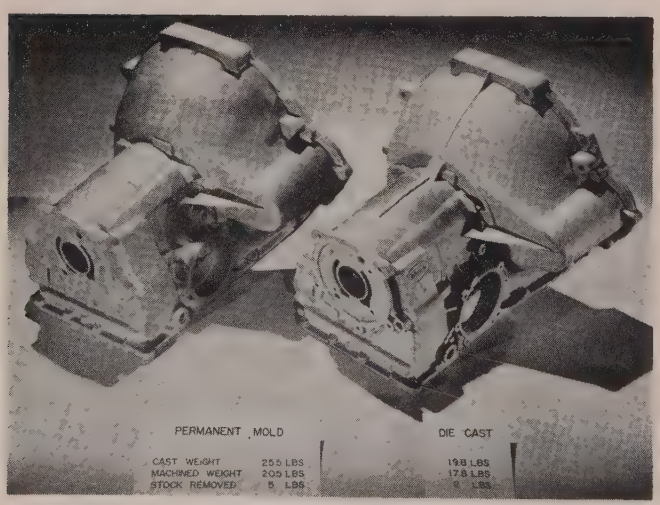


Fig. 16 — Transmission housing

Since gray iron parts are normally cast at the most suitable wall thickness, we can often reduce the walls in our efforts to design for uniform stresses in all sections through stress analysis which creates the most efficient strength-weight ratio. Thinner walls in combination with cored holes, the elimination of core shift allowances and the reduction of machining stock greatly decrease the weight of aluminum die castings far beyond the density ratio of 2.5/1 between gray iron and aluminum. If full advantage of the many weight-saving features of aluminum dies casting is taken, this weight ratio can be boosted to an economical 4/1 as you see in the three examples of this chart. An engine block which weighs 208 lb in unmachined gray iron can be converted into an aluminum die cast block of ony 55 lb (round figures — 200 lb/50 lb).

There is no other casting process in which the weight of a part can be held as low as in aluminum die casting. In comparison with aluminum permanent mold, die castings of equal strength are in average 20% lighter. Fig. 16 demonstrates this weight advantage of die casting over aluminum permanent mold casting on a transmission housing which is now in production. (Since engine blocks in aluminum permanent mold are not yet in existence, I had to refer to this interesting example.)

The reduction in the as-cast weight between 25.5 lb in permanent mold and 19.8 lb in die casting is a direct saving of 5.7 lb of aluminum. The difference between the weight of the machined parts shows that the die cast housing is almost 3 lb lighter. On the permanent mold casting, 5 lb of chips have to be machined off; on the die casting, only 2 lb — or, a saving in machining of 60%.

This transmission housing has been tooled up for aluminum permanent mold because it was felt that heavy sections in this housing may present a porosity problem in die casting. But recent tests with these housings have proved that die castings can be at least as sound as permanent mold castings, and because of their higher fatigue strength are even stronger in spite of finely dispersed microporosity in the center of the die cast sections.

Hypereutectic Al-Si Alloys

The conversion problems described may show you that the introduction of aluminum die castings for engine blocks necessitates a complete redesign of the engine. It actually means to adapt the engine block to the mechanical and physical properties of aluminum in order to get the most efficient design at the most economical level. *Any compromise in the hope to get away with less changes will result in an inferior design from either a technical or cost standpoint.*

Since we engineers are by nature always reluctant to changes, the rediscovery of the hypereutectic Al-Si alloys (Table 2) was greeted with enthusiasm by many an automotive engineer in the hope that unprotected cylinder bores in these alloys will give satisfactory wear results and that the basic changes with cast-in gray iron sleeves described above can be avoided. I said "rediscovery" because these hypereutectic Al-Si alloys have been used in Europe for many, many years. Their discovery was patented in Germany as early as 1924. Ever since, they have been used in Europe for parts where low

thermal expansion was of importance. But since they are not easy to cast and difficult to machine, there is even today only one field in which these hypereutectic Al-Si alloys could establish themselves, and that is for pistons of 2-stroke engines and diesel engines. Today, several million pistons per year, or approximately 20% of the German piston consumption, are made in hypereutectic Al-Si alloys, using the permanent mold process.

Since these alloys are not used to any great extent in any other casting method nor for any other casting field and since no complete data on their properties are available in this country, the values in the chart refer to European data on permanent mold pistons except the column on the extreme right which shows comparative data on gray iron, and the column on the extreme left which gives the U.S.A. data of a standard die casting alloy of the eutectic type with 12% Si.

The most outstanding property of these hypereutectic Al-Si alloys certainly is their low thermal expansion which decreases from 11.5 to 9.2 as the Si content increases. On the other hand, the thermal conductivity is somewhat impaired with higher Si content and falls off from 0.29 to 0.22. It is claimed that these alloys have somewhat higher strength and hardness at elevated temperatures. However, as you see from the table in the temperature range up to 300 F to which engine blocks may be subjected, the difference is hardly noticeable.

Properties which are detrimental to the introduction of these hypereutectic Al-Si alloys for engine blocks are, among others, their extremely low elongation which makes the parts quite brittle. Extremely high melting and even higher casting temperatures — well above the liquidus point — have an adverse effect on the manufacturing cost. These higher casting temperatures — some 200 deg above our standard alloys — increase the danger of microporosity, volume-shrinkage and segregation. In our die casting tests we found, in addition, a hot shortness and excessive cracking in sharp corners of the casting, and a porosity problem which is more severe than in standard die cast alloys. The hypereutectic alloys, therefore, are difficult to cast in sand and in permanent mold, but more difficult to produce in die casting.

Also, the extremely high casting temperatures badly affect the lifetime of molds and dies. In permanent mold it is reported that the mold life is reduced by 20% in comparison with a standard aluminum alloy. Since not enough experience in die casting is available, their effect on die casting dies is not shown in the chart. However, we must assume that the life of a die casting die will be affected somewhat more than a permanent mold die, maybe by some 25–30%.

As the last line of Table 2 shows, the machinability is also adversely affected by increased Si in the alloy. The free Si crystals are extremely hard and, therefore, wear out the cutting tools twice as fast at an 18% Si content. Alloys with 24% Si wear out the tool bits four times as fast as a eutectic Al-Si alloy. In addition, the cutting speed has to be reduced appreciably with increased Si content.

These difficulties in casting and machining have

Table 2 — Physical Data on Hypereutectic Al-Si Alloys in Comparison to Standard Al-Alloys and Gray Iron

Class of Alloy	U. S. Data Die Casting	European Data on Permanent Mold Pistons				Gray Iron
		Eutectic		Hypereutectic		
Basic Materials	Aluminum	Aluminum	Aluminum	Aluminum	Aluminum	Iron
Main Additions	12% Si	12% Si	18% Si	21% Si	24% Si	SAE 120
Secondary Additions	None	1% Cu, 1% Ni, 1% Mg		1% Cu, 1% Ni, 1% Mg		
Thermal Expansion, in./in./F × 10^{-6}	11.5	11.4	10.3	9.7	9.2	6.1
Thermal Conductivity, cal/cm/sec/C	0.29	0.31	0.28	0.27	0.25	0.12
Brinell Hardness, kg/mm²						
At 68 F	80–110	90–125	90–140	90–120	90–125	200
At 300 F		75– 95	80–110	80–105	70–105	200
At 480 F		35– 50	40– 70	50– 70	45– 65	170
Ultimate Tensile Strength, psi			29,000	28,500	28,500	35,000
At 68 F	39,000	31,000	24,000	24,000	23,500	
At 300 F	27,000	25,500				
At 480 F	14,500	14,200	15,000	15,000	15,000	29,000
Elongation, %			0.2–0.8	0.2–0.5	0.1–0.3	0.9
At 68 F	0.8–1.5	0.3–1.0	1.3–2.3	1.2–2.1	1.0–1.8	
At 480 F	3.0–5.0	3.0–5.0				
Modulus of Elasticity,			11.4	12.3	12.3	15.0
psi × 10^{-6}	10.3	10.7	0.9	0.8	0.7	0.4
Relative Abrasion Value	1.0	1.0				
Casting Temperature, F	1250	1300	1400–1430	1480–1520	1570–1610	
Reduction of Mold Life Compared to Eutectic Al-Si Alloys			10%	15%	20%	
Increase in Cutting Tool Wear Compared to Eutectic Al-Si Alloys			Twice	3 times	4 times	

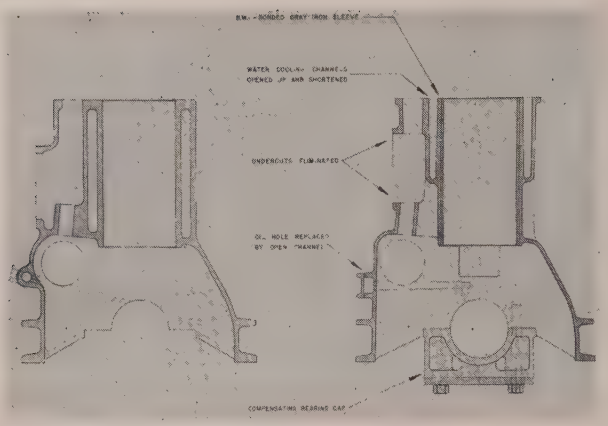


Fig. 17 — Six-cyl engine block in gray iron converted to aluminum die casting

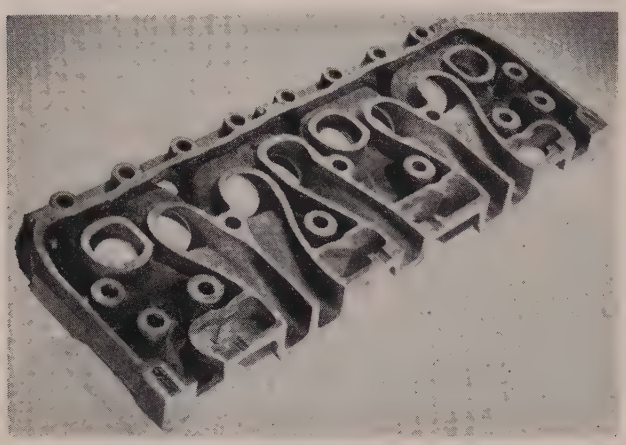


Fig. 18 — Cylinder head (cut through gray iron)

limited the use of these alloys to thermally highly stressed pistons. Since no other castings have been made in these alloys, it is, in my opinion, a far stretch of imanigation to hope that complete engine blocks can be made in these hypereutectic alloys at an economical level.

From a technical standpoint, it has to be said that the first satisfactory wear results which have been obtained in laboratories may be somewhat misleading. Europeans, who have so much experience in the manufacture of pistons in these hypereutectic Al-Si alloys, have also tried to use them for cylinder sleeves and cylinder barrels. All their tests, however, showed that none of the aluminum alloys tested, including the hypereutectic group, is 100% foolproof. While very satisfactory results were obtained under normal operating conditions, failures occurred with excess dirt in the oil and especially during cold starts, or at low oil levels in the crankcase, because the aluminum alloys generally do not have the same capacity to retain oil as gray iron does.

Field tests made in Europe actually confirm Alcoa's findings that under ideal conditions uncoated sleeves in hypereutectic Al-Si alloys show less wear than gray iron sleeves. Since, however, the situation may suddenly change when marginal lubrication prevails, all European experts I contacted advised that unprotected cylinder bores in hypereutectic Al-Si alloys are, in their opinion, not suitable for use in passenger-car engines. I believe that when the first storm of novelty has faded and has given way to the cold facts, the American automotive industry will also come to the conclusion that it cannot jeopardize its high reputation of dependability by only one failure in 10,000 engines.

Future Engine Block Design

In summarizing the development of a die cast engine block for water-cooled engines, we show in Fig. 17, a typical cross-section of a 6-cyl gray iron block and its proper redesign in aluminum die casting containing the basic features mentioned in this paper.

Taking in consideration all the properties of the aluminum alloys available, we believe that the best choice for engine blocks is a eutectic Al-Si alloy with 12% Si known under ASTM Standard as S12A. This alloy has excellent castability and good mechanical properties, is not too difficult to machine, and produces sounder castings than any other aluminum alloy. In the past it has always been used for extremely intricate castings (used exclusively for valve bodies). It has the lowest thermal expansion of all standard casting alloys.

The elimination of undercuts is demonstrated in many ways, and the dry sleeve arrangement with B.M.I. bonded gray iron sleeves has been chosen at this stage of development as the most economical and already proved design. The higher thermal conductivity of aluminum is utilized by reducing the water channels to cover only that portion of the bore which is exposed to the hot combustion gases. Walls chosen are in average only 5/32 and heavy sections have been avoided. Since it would have been difficult to cast the oil supply hole through the whole length of the block, it has been changed to an open channel closed up by a simple cover. An expansion compensating bearing cap prevents excess clearance in main bearings when the engine heats up. These basic design features demonstrated on a 6-cyl in-line engine block can also be applied to a V-8 engine block.

Cylinder Head and Manifold for Water-Cooled Engines

The conversion to an all-aluminum engine is complete only when the major components of the engine block, its cylinder head, and its intake manifold also have been changed to aluminum. While the redesign of the water-cooled engine block has found many satisfactory solutions, these two components still present serious conversion problems because of their many internal channels and pockets now made with disposable sand cores.

Fig. 18 is a cut through the intricate channel system of a cylinder head with its many pockets for the cooling water, the channels for the incoming gases, and for the exhaust gases. Many attempts have been made to eliminate the undercuts by completely redesigning these parts. However, these channels are so restricted in their location that no economically acceptable solution for a *one-piece* design in die casting has been found yet.

The present design of cylinder heads, therefore, lends itself only to sand casting or semi-permanent

mold, but both these processes are very expensive because of the high costs for the sand cores.

A more acceptable solution (Fig. 19), and probably the most economical approach, is to split the cylinder head into a main body and a cover with water channels open to both sides, thus eliminating the undercuts. The two parts may be joined with a plastic adhesive into the assembled piece.

Tests made with a number of plastic adhesives indicate very satisfactory physical properties. The plastic bond shows good corrosion resistance, a tensile strength between 6000 and 10,000 psi at room temperature, and a resistance to heat up to 500 F.

It must be admitted that the cure time and cure pressure to guarantee a good bond still represent some problems which are not completely solved yet. However, since smooth passages, the elimination of sand pockets, and uncontrollable core shift are important advantages over sand cast cylinder heads, the split design has found much interest and several different solutions are under consideration.

Intake Manifold — The internal channel system of an intake manifold is just as complicated as that of the cylinder head. Therefore, the only solution to a die cast manifold is a split design similar to the cylinder head. An economical design is possible if the parts can be joined without machining the joining areas and fitting them together with a tongue and groove arrangement. Tests with unmachined joints have successfully been concluded and manifolds in a split die cast design may already be used in 1960.

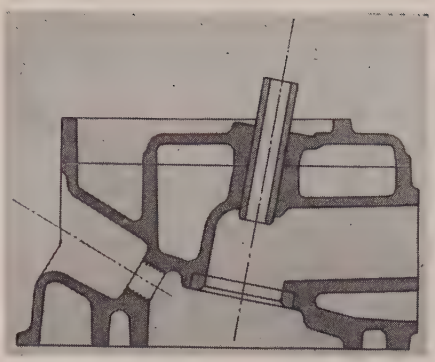


Fig. 19 — Cylinder head — split design (die casting)

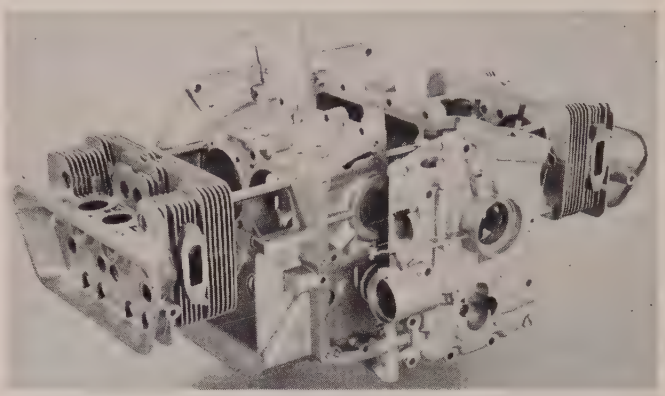


Fig. 20 — Aircooled engine (Porsche)

Aircooled Engines

With an eye on the smaller, more economical European cars, aircooled engines are also under consideration. Fig. 20 shows the major aluminum castings of the aircooled 4-cyl Porsche engine with a total weight of 70 lb.

The cylinder barrels are not shown because they are not made as aluminum castings. For their more expensive models, Porsche buys these barrels machined from extruded aluminum with the bore protected by a porous chrome plate. In the standard Porsche models, the cylinder barrels were changed back to gray iron to reduce costs.

This engine is shown because its basic features of opposed cylinders, of the split design of crankcase and transmission case, and the arrangement of the engine in the rear are also the basic design features of the Volkswagen engine, which have in recent months been adopted by companies in this country. The Volkswagen engine is very similar to the Porsche engine, but the main difference is that all its major castings for cankcase and transmission are made in magnesium die casting except the cylinder barrels which are gray iron, and the cylinder head which is aluminum permanent mold.

All the castings of the Porsche engine are produced by the permanent mold process because of their low production rate. The parts of the new aircooled 6-cyl Chevrolet engine will be produced by the low-pressure casting method, a somewhat improved and highly mechanized variation of the permanent mold process. Yet for similar engines of the aircooled type, most of these castings have been designed for aluminum die casting, even including the cylinder head, which together with transplant-coated cylinder barrels in aluminum die casting will result in a more economical design and a more efficient engine.

Conclusion

With these examples I tried to convey to you the important role the die casting technique has played in the past and will play in the future development of aluminum engines. Some of the conversion problems have been stated and solutions for the water-cooled engine block described in details. It was, however, pointed out that there are still many problems to be solved on such important components as the cylinder head and manifold.

Without revealing any secrets, it can be said that the first die cast water-cooled 6-cyl engine blocks for a passenger-car engine made their debut in 1959, but production on a larger scale did not get under way before 1960. Several hundred cars equipped with die cast 6-cyl aluminum engine blocks have successfully passed the first stage of severe field tests.

When I presented my first paper on engine blocks several years ago, I predicted that die cast aluminum engine blocks will be introduced in 1960 or 1961. Judging by the little progress made at the beginning, I felt that I was badly off with my crystal ball prediction. However, the interest in aluminum engines has suddenly picked up momentum and now it seems that my prediction may come *true after all.*

preliminary design

Structure of a

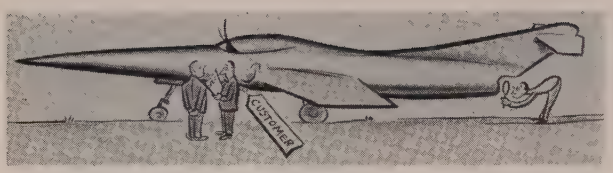


Fig. 1 — Requirements: loads, stability and control, safety and reliability

Fig. 2 — Shape: thin and long, flexible, heavy

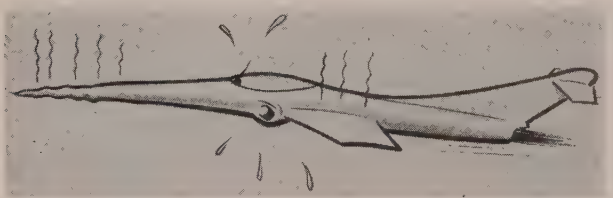


Fig. 3 — Heating: new materials, insulation, buckling

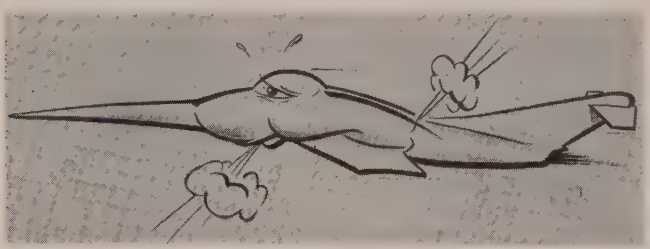


Fig. 4 — Altitude: leakproof, fail proof, goof proof

THE PAST two decades have not seen a substantial change in the basic structure of transport-type aircraft. The heritage of the Boeing 247, Douglas DC-2, and the original Lockheed Electra is clearly discernable in their grandchildren the Boeing 707, the Douglas DC-8, and the latest Lockheed Electra. There has even been a considerable amount of cross-breeding and, except for certain details, an unitiated layman might not be able to identify the structure with the company producing it. It is true there has been a steady evolution which has taken advantage of higher strength materials and improved production methods; but in general, the structural configuration of airframes for transports has become rather standardized.

The structures engineer has been busy improving his efficiency in methods of analysis, taking advantage of modern aids such as high-speed computers. He has also been working on the structure itself to improve its efficiency and reduce the cost. Certain problems have come to the foreground such as fatigue and fail-safe and he has tackled them with vigor to obtain solutions. In recent months, the advances in technology which have made possible an economic supersonic transport has provided the structures engineer with a new challenge — namely, to provide a structure operating under more stringent environmental conditions than heretofore, but which nevertheless is light in weight and economical to build.

The purpose of this paper is to discuss some of the problems which face the structures engineer as he meets the challenge and to suggest some possible solutions to these problems.

Problem Areas

The main problem areas which are directly related to the structural configuration of a supersonic airplane, whether it is a transport or some other type can be categorized in the following way:

1. What are the requirements for such an air-

considerations for the

Trisonic Transport

M. G. Childers
Lockheed Aircraft Corp

plane? (See Fig. 1). Among the many things to be considered are the loads and their distribution to the airframe, stability and control requirements, and the safety and reliability standards. Current regulations, which apply in the main to subsonic transports, do not adequately cover the supersonic transport and a thorough analysis of the requirements and criteria is necessary if a successful supersonic transport is to be built. A few papers have been published[1,2,3,4,5] and additional work is being done on this subject.

2. In order to achieve the high speeds being considered with reasonable powers and airplane size, the drag must be low, which leads to shapes that are very thin (Fig. 2). The problems which result are obvious and considerable weight can result from the necessity for making the structure stiff enough to provide adequate stability and control and to prevent flutter. Shape distortion also may lead to increased drag and decreased performance.

3. The high speed results in aerodynamic heating, which in turn opens up a variety of new problems for commercial transports with which the structures engineer has not been particularly concerned in the past (Fig. 3). Scores of excellent articles have been written in the past few years on the subject of aerodynamic heating and its effects on the structure.[6,7,8,9,10,11] This paper will treat the subject of aerodynamic heating in relation to the other problems and as a facet of the overall design problem.

4. In order to fly economically at these speeds, it is necessary to fly at altitudes which are roughly twice those now being used for subsonic jet transports (Fig. 4). This means that problems related to passenger comfort and structural reliability are more critical than they have been in the past. For example, leakage rates must be kept to very low values to reduce the amount of power required to keep the cabin pressurized, and fatigue life and fail-safe considerations are more important than they ever have been.

Structural Parameters

Some of the important parameters to be considered in the selection of a structure for a Mach 3 transport can be placed under three categories as shown in Table 1. In any design which is intended to be operated at a profit by the airlines, the economic parameters such as initial cost and service utilization become very important and will weight heavily in the selection of structure. The structural parameters and economic parameters are not always compatible and optimization studies are required to evaluate the best compromises.

Value of Pound

If weight is selected as the common denominator for the evaluation of these parameters, the question is asked "What is the effect of using a more expensive structure to reduce the weight of the structure?" In other words, "What is the value of a pound of

---

[1] "Design Criteria for Heated Aircraft Structures," by Robert Goldin. Paper presented at ASME Symposium on Structures for Thermal Flight, March, 1956. (Paper No. 56-AV-14)

[2] "Discussion on Safety Factor Requirements for Supersonic Aircraft Structures," by George M. Goldman. Paper presented at ASME Symposium on Structures for Thermal Flight, March, 1956. (Paper 56-AV-18)

[3] "High Temperature Flight Safety Factors," by A. T. Clemen. Paper presented at Third Symposium High-Speed Aerodynamics and Structures, San Diego, March, 1958.

[4] "Thermal Creep Design Criteria," by Robert Goldin. Paper presented at IAS National Summer Meeting, Los Angeles, June, 1957.

[5] "Aircraft Structural Factor of Safety," by George N. Mangurian. AGARD Report 154, presented at Sixth Meeting of Structures and Materials Panel, Paris, November, 1957.

[6] "High Temperature Effects in Aircraft Structures," by N. S. Hoff, ed. Pub. by Pergamon Press, New York, 1958.

[7] Symposium on Structures for Thermal Flight, ASME, Los Angeles, March, 1956.

[8] "Structural Design for Aerodynamic Heating." WADC TR 55-305, October, 1955.

[9] "Thermal Stresses," by B. E. Gatewood. Pub. by McGraw-Hill Book Co., Inc., 1957.

[10] "Effects of Physical Factors and Analytical Procedures on Predicted Temperatures and Thermal Stresses," by R. E. Hubka. L.R. 12777, Lockheed Aircraft Corp., California Division, December, 1958.

[11] "Thermal Stress Analysis for Aircraft Structures," by B. A. Boley and J. H. Weiner. WADC TR 56-102, August, 1955.

STRUCTURAL MATERIALS for Mach 3 jet transports pose difficult problems for the design engineer. Reasons for this problem are the incomplete information available on the many possible metals and the diversity of critical properties that are added by supersonic requirements.

The material properties discussed in this paper include tensile strength, resistance to crack propagation, ease of fabrication, weldability, and thermal expansion. Cost factors are also considered.

The structural configuration of the wing and fuselage is an example of the complexity of the material selection problem. The wing may be rigidity-critical, and the fuselage strength-critical; each requires diferent material properties to solve the problem.*

weight?" This, of course, is a controversial question and no attempt will be made here to settle it. However, by making some rather simple assumptions it is possible to arrive at a figure based on the direct operation cost formula which gives a feel for at least one value.

The following assumptions are made:

1. The upper limit of allowable cost is where the direct operating cost (DOC) of the reduced weight airplane equals the DOC of the comparison airplane.
2. The speed, range, payload, and other performance characteristics remain unchanged.
3. Only the cost of the airframe structure is affected.
4. The airplane size is adjusted to take advantage of the decreased weight.

The standard ATA direct operating cost equation can be put in the form:

$$\text{DOC} = K_1 + \frac{K_2}{V_b} + \frac{K_3}{V_b} + \frac{K_4 F}{D_b} \qquad (1)$$

where:

K_1 = Fixed cost, \$/mile
K_2 = Cost factor of airplane elements unaffected by structure. Under assumption (3) above this factor is constant, \$/hr
K_3 = Cost factor for airframe structure, \$/hr
K_4 = Unit fuel cost \$/lb
V_b = Block speed, mile/hr
D_b = Block distance, mile
F = Fuel weight, lb required for D_b

Combining the first two terms, which under the assumptions are both constant, the basic airplane DOC will be

$$\text{DOC} = K_5 + \frac{K_3}{V_b} + \frac{K_4 F}{D_b} \qquad (2)$$

The lighter but more expensive airplane will have a DOC of:

$$(DOC)_1 = K_5 + \left[\frac{K_3 + \Delta K_3}{V_b}\right] + \left[\frac{K_4 (F - \Delta F)}{D_b}\right] \qquad (3)$$

where:

ΔF_3 = Increment of increased cost due to decreased weight
ΔF = Increment of reduced fuel which results when the weight and size are decreased

From assumption 1:

$$\Delta K_3 = \frac{K_4 \Delta F V_b}{D_b} \qquad (4)$$

To relate this to weight and the cost per pound of weight saved:

$$\Delta K_3 = \frac{X \Delta W}{T} \qquad (5)$$

where:

X = Cost per pound of weight saved
ΔW = Weight saved
T = Total flight hours
$\Delta F = Y \Delta W$

$$Y = \frac{\Delta F}{\Delta W} = \text{Fuel saved per pound of structural weight saved} \qquad (6)$$

Therefore, from Eqs. 4–6:

$$X = K_4 \times Y \times \frac{V_b T}{D_b} \qquad (7)$$

$$X = K_4 \, YN \qquad (8)$$

$$N = \frac{V_b T}{D_b} = \text{Number of flights in life of aircraft}$$

This surprisingly simple result states that the maximum allowable cost for saving a pound of weight is a function of the unit cost of fuel times the number of anticipated flights in the lifetime of the airplane modified by a factor which relates the fuel saved to the weight saved. The relation plots as in Fig. 5.

As an example, suppose that $Y = 1$, the cost of fuel is \$0.028 per lb, and the anticipated number of flights of the airplane is 12,500 (corresponding to 10 years of life at 8 hr a day utilization over 3000-mile range at 1280 miles per hr block speed). The maximum allowable cost then becomes \$350 per lb. For a piston-engine transport using the same set of conditions, except for speed, the value of a pound of weight saved from this formula is about \$85. The validity of this method can be argued and various factors and modifications can be made for specific applications, but the ratio between the two is essentially correct and so it is concluded that a pound of weight saved in a Mach 3 transport is worth four times the amount considered reasonable for a piston-engine transport. This is a very significant input into the whole system as various compromises are considered.

Weight Ratio

A common equation for evaluating the efficiency of a given airplane is the Breguet range equation:

$$R = K \left(\frac{L}{D}\right)\left(\frac{M}{SFC}\right)\left(\log \frac{W_1}{W_2}\right) \qquad (9)$$

The first portion of the equation, that is, $\left(\frac{L}{D}\right)\left(\frac{M}{SFC}\right)$

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 9, 1959.

relates to the aerodynamic and propulsion efficiency and was discussed by Hibbard and Bailey.[12] The last term in the equation, W_1/W_2, generally known as the weight ratio, is sometimes referred to as an index of structural efficiency.

In order that a common basis for the following discussion is established the weight ratio is defined as follows. (These definitions are not strictly applicable in the pure sense to the Brequet equation.)

W_1 = Airplane take-off weight

W_2 = Airplane weight at landing after traversing the Range R

W_2 is thus W_1 – Trip fuel. It is seen, therefore, that

$$W_2 = W_A + W_P + W_S + W_L + W_E + W_{RF}$$

where:

W_A = Weight of the airframe (structural weight)

W_P = Weight of the propulsion system

W_S = Weight of the operating systems

W_L = Weight of the payload

W_E = Weight of the equipment, crew, and the rest

W_{RF} = Weight of reserve fuel (zero for max range)

It is readily seen that the weight ratio is not a reliable index of structural efficiency since many things enter into its determination which are not directly related to the structure. Typically, W_A may be only 50–55% of W_2. This means that a decrease in structural weight does not have a one-to-one influence on the weight ratio.

To illustrate the role of the weight ratio further, a survey was made of most of the multiengine airplanes in current usage to correlate the weight ratio with take-off weight. The results are plotted on Fig. 6. For convenience, since most of the readily available information was in terms of take-off weight and weight empty instead of zero fuel weight, the data on Fig. 6 are plotted in terms of:

(Gross take-off weight)/(Weight empty).

This parameter has an additional advantage in that for some classes of airplanes it is difficult to distinguished between the payload and the fuel load.

Some interesting observations can be made about the information contained on this chart:

1. The prop-jet airplanes and jet airplanes lie in the same scatter band which overlaps but is generally higher than the piston-engine airplanes. This does not mean that the structural efficiency is better but comes about because of conditions not directly related to structure.

2. The bomber-tanker scatter band is highest of all. The fallacy of judging the structural efficiency on the basis of weight ratio is apparent when two similar airplanes are compared — one a tanker and the other a transport. The tanker lies in the upper region of the upper scatter band, the transport lies in the jet transport scatter band.

3. As a first approximation it is assumed that a supersonic transport must lie within the jet transport region if it is to be competitive. This means that the systems as well as the structure must be carefully designed to obtain the lightest weight commensurate with cost. From this curve then, assuming a supersonic transport has a take-off weight of 250,000 lb, the weight empty would lie between 109,000 and 119,000 lb, and assuming the structural weight is 55% of the weight empty, the structure should weigh 60,000–65,000 lb.

Materials

The discussion thus far has been concerned with establishing some target numbers which can be used to judge the suitability of the various structural configurations which might be proposed. Before going on to a discussion of such configurations, however, one other very important topic must be treated — the materials to be used in the airplane.

It is obvious that since the major portion of the structure operates at temperatures above 350 F, aluminum alloy, our old standby, is no longer a suitable material. Aluminum must, therefore, be supplanted by other materials which will withstand the temperatures. At first there appears to be a variety of candidates. Table 2 is a partial list of such possibilities. In addition to this list, which represents only a few of the possibilities, other alloys appear on the market at frequent intervals, which add to the confusion. Furthermore, as more is learned about the materials and as techniques for manufacturing and fabrication are developed,

Table 1 — Structural Design Parameters Supersonic Transport

Structural Efficiency	Initial Cost	Service Utilization
Material density	Material cost	Ruggedness
Allowable stress	Fabrication ease	Repairability
Stiffness	Tolerances	Inspectability
Smoothness	Tooling	Reliability
Fatigue resistance	Plant facilities	Accessibility
Fail safe	Inspection methods	Material stocking
Thermal environment	Time span	Base facilities
External shape	Health precautions	Interchangability

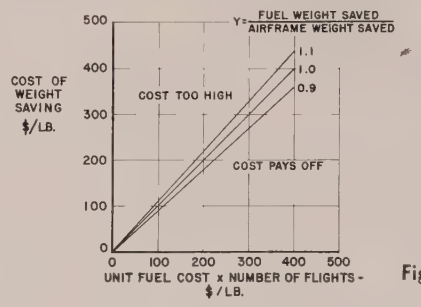


Fig. 5 — Value of pound of weight saved

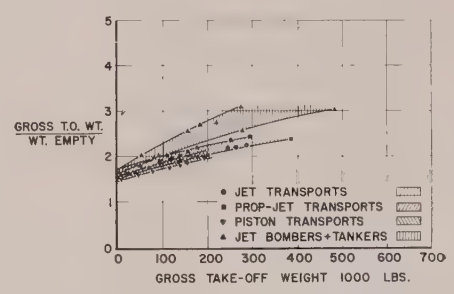


Fig. 6 — Weight ratio = gross take-off weight / weight empty versus gross take-off weight

[12] "Case for Supersonic Transport," by H. L. Hibbard and R. A. Bailey. IAS paper No. 59-119.

the choice is subject to change.

The choice is further complicated by the fact that not all the important properties are available for a specific material and in many instances the properties which are available do not agree between sources. A list of properties which must be known in order that a design job can be done is shown in Table 3. It is usually possible to find room temperature ultimate and yield tensile strengths and elongation, at least typical values, some of the physical properties such as density and coefficient of thermal expansion and heat-treat cycles. Often it is possible to obtain some of the temperature effects on the properties, particularly if only the short time effects are wanted. However, data on crack propagation, or compressive stress-strain curves, or effects of long-time aging at operating temperatures are generally lacking. Of course, more data of this sort are becoming available as time goes on and the material suppliers, government agencies, and fabricators are working toward obtaining the necessary information. However, it is a long, tedious, and expensive process and it will be some time before the designer will be able to pick up a handbook much as ANC-5 and find the information he needs on a specific material.

The designer must, therefore, base his material selection on the best information he is able to garner, incomplete though it might be. He must also be a prophet and try to project his consideration into the future possibilities. For example, he must not overlook a possible material of the future such as beryllium which, due to its high ratio of modulus to density, potentially could reduce the structural weight by a substantial amount. (See Hoffman.[13]) For the present, however, there are some serious deterrents to its use and further development is required before it can be considered as a suitable structural material for a commercial transport.

No attempt will be made here to discuss in detail all the various properties listed in Table 3. Rather a few typical properties in each category will be touched on briefly to indicate the problem facing the designer. The materials chosen for investigation are the titanium alloys and stainless steel.

Table 2 — Materials for 400-800 F Usage

Category	Material Designation	Remarks	
Magnesium	Magnesium-Thorium alloy	Low strength Short time at temperature	
Aluminum	Aluminum powder metals	Low strength Secondary structure only	
Titanium	6A1-4V	Available	Expensive
	4A1-3 Mo-IV	High strength	
	120 VCA beta	Unstable above 500 F	
	Commercially pure	Low strength	
Corrosion resistant steels	301 extra full hard	Fatigue & fail safe questionable	
	AM 350	Good fail safe	
	17-7 PH	High strength	
	15-7 Mo	High strength	
	420	Fatigue & fail safe questionable	
Low-alloy steels	SAE 4350	Poor corrosion resistance	
	SAE 4130	Better for higher temperatuers	
	SAE designation H11	Good for welded steel tubing	
Nickel base	Inconel X	Too heavy for this temp. range	
	Rene 41	Better for higher temperatures	
	Hastelloy	Expensive	
Miscellaneous	Beryllium	Expensive	
	Cobalt base alloys	Difficult to fabricate	
	Molybdenum	Better for higher temperature	
	Vanadium	Heavy (except beryllium)	

Tensile Properties — Most comparisons which are made between the various materials are on the basis of some mechanical property such as ultimate strength. Fig. 7 is typical of such data for various titanium and stainless-steel alloys. When considered on this basis alone, the titanium alloys appear to be superior. This type of presentation can be expanded and much useful information can be summarized in the form of various criteria plotted against temperature. An excellent presentation of this sort has been made by Ekvall[14] who discussed various materials in terms of tensile ultimates, tensile yield, stiffness, buckling, and crippling.

Crack Propagation Properties — Another property of materials has emerged in recent years and is of major importance — that is, its resistance to crack propagation. This information is necessary in the evaluation of the fail-safe characteristics of the structure. As various materials are evaluated, it is discovered that some of the materials which appear to be good from a strength-weight standpoint fall down miserably when its crack propagation resistance is determined. Obtaining the data is expensive and consequently not much information is available on these alloys, particularly at temperatures above room temperature. Furthermore, the interpretation of the data is subject to question since test methods are not standardized. Some of the data available to the author on the materials of interest are presented in Figs. 8 and 9. Fig. 8 presents data on three stainless steels which are candidates for use in the airplane. Note the lack of consistency in behavior between the various steels, and that in general the transverse properties are only about one-half as good as the longitudinal properties for the better alloys.

Fig. 9 shows some typical data on titanium alloys and compares these alloys with 2024-T3 aluminum alloy and AM 350 SCT stainless steel. Here again the transverse properties could stand improvement. More information is required to establish design information for the titanium alloys. These data will strongly influence the choice of materials to be used for the pressurized fuselage, for example, keeping in mind the necessity for providing a fail-safe structure.

Fabrication — The fabricability of the material, while not directly a structural property, is of vital interest to the structure designer since it influences the type of structure he can design. For example, small bend radii are always an advantage structurally and sometimes, where space is a limitation, may be a necessity. It is a fact that the materials which must be used for these airplanes are more difficult to fabricate than the aluminum alloy. The structural design must, therefore, cater to the properties which are peculiar to the material so as to make the production task as easy as possible. In addition, of course, the shop must develop its tech-

[13] "Beryllium as Structural Material," by G. A. Hoffman. *Aeronautical Engineering Review*, Vol. 16, February, 1957, pp. 50–55, 82.

[14] "Strength and Thermal Limitation of Materials for Airframe Components," by J. C. Ekvall. SAE Transactions, Vol. 67, 1959, pp. 590–595.

niques and equipment to handle the material and reduce the cost. The situation is analagous to that which existed years ago when the industry shifted from wooden structures to all metal structures — the pot of glue, bag of nails, and the plane and sandpaper being replaced by the rivet gun, power brake, and hydropress. For the new generation of airplanes these will be supplemented by the high-speed fully automatic spot welder and fusion welder, and the brazing furnace.

Table 4 shows the results of an evaluation of the fabricability of the six alloys chosen for discussion. It should be remembered that at this point in the "State of the Art," considerable speculation is inherent in a listing of this kind. From this chart it appears that the B-120 VCA material is the best from a fabricability standpoint. However, it may be limited in its application for a flight vehicle which is designed to operate at Mach 3 or better for long periods of time, since it appears to be metallurgically unstable at temperature.

Weldability — Another important characteristic which must be considered is whether the material can be spot welded, fusion welded, or brazed. Although aluminum airplanes can be spot welded and indeed some are, the ease with which holes can be drilled and rivets formed made extensive use of spot welding less attractive. Fusion welding of the structural aluminum alloys is, of course, unacceptable in any case. For the titanium and stainless-steel alloys the situation is completely different. The difficulty of drilling holes and forming rivets make extensive use of this time honored method of assembly less desirable, and it is necessary that the material can be at least resistance spot welded. In addition, it is highly desirable that satisfactory fusion welds can be made and that the material can be brazed. Generally speaking, the stainless steels are more readliy welded and brazed than are the titanium alloys. However, satisfactory spot welds can be made in titanium, although at the present time fusion welds and brazed joints in high-strength titanium are questionable. Further effort in this field may result in a satisfactory solution to the problem.

Thermal Properties — The most important thermal property is the coefficient of thermal expansion. Thermal conductivity, specific heat, and emissivity may be important for specific problems, but generally speaking large variations in these properties result in relatively small differences in the thermal stresses.

Without going into detail, since the methods of calculating the thermal stresses has been discussed by others,[6,7,8,9,10,11] the nature of the problem is such that high values of thermal conductivity and emissivity and low values of specific heat result in a more even temperature distribution in the struc-

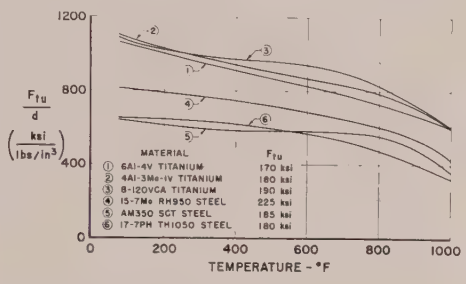


Fig. 7 — Ultimate tensile strength-density ratio versus temperature

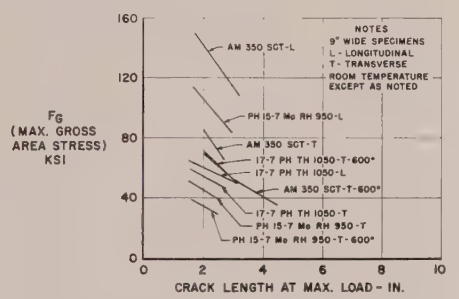

NOTES
9" WIDE SPECIMENS
L - LONGITUDINAL
T - TRANSVERSE
ROOM TEMPERATURE EXCEPT AS NOTED
AM 350 SCT-L
PH 15-7 Mo RH 950-L
AM 350 SCT-T
17-7 PH TH 1050-T-600°
17-7 PH TH 1050-L
AM 350 SCT-T-600°
17-7 PH TH 1050-T
PH 15-7 Mo RH 950-T
PH 15-7 Mo RH 950-T-600°
F_G (MAX. GROSS AREA STRESS) KSI
CRACK LENGTH AT MAX. LOAD - IN.

Fig. 8 — Crack tear test results, stainless steels

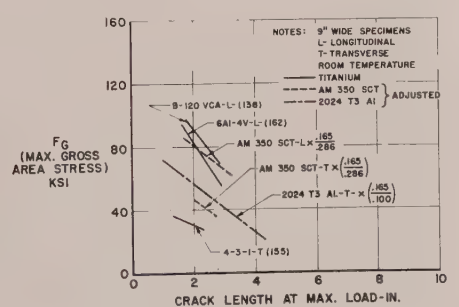


Fig. 9 — Crack tear test results, titanium alloys

Table 3 — Required for Material Evaluation

Mechanical	Physical	Metallurgical	Miscellaneous
1. Tensile properties	1. Density	1. Chemical composition	1. Cost
2. Compression properties	2. Thermal conductivity	2. Heat-treatment	2. Availability
3. Shear properties	3. Thermal expansion	3. Critical temperatures	3. Toxicity
4. Bearing strength	4. Specific heat	4. Aging effects	4. Tolerance
5. Poissons ratio	5. Emissivity	5. Spot welding	
6. Impact strength	6. Electrical resistance	6. Fusion welding	
7. Fatigue properties	7. Magnetic permeability	7. Brazing	
8. Creep in tension & compression	8. Nuclear properties	8. Corrosion & oxidation	
9. Stress rupture	9. Melting temperature	9. Stability	
10. Notch sensivitity			
11. Formability			
12. Machinability			
13. Crack propagation			
14. Hardness			
15. Thermal shock			

Table 4 — Ease of Fabrication

Material	Forming	Heat-Treat	Machinability
Stainless Steel 15-7 MO AM 350 17-7 PH	Can be formed readily at room temperature in the annealed condition	Required for strength — complicated and subject to warpage	Fair
Titanium 6A1-4V 4-3-1	Hot forming for all but simplest bends	Water quench-subject to warpage	Difficult
B-120 VCA	Can be formed at room temperature or at moderate temperatures	Simple agign to achieve strength (may overage)	Difficult

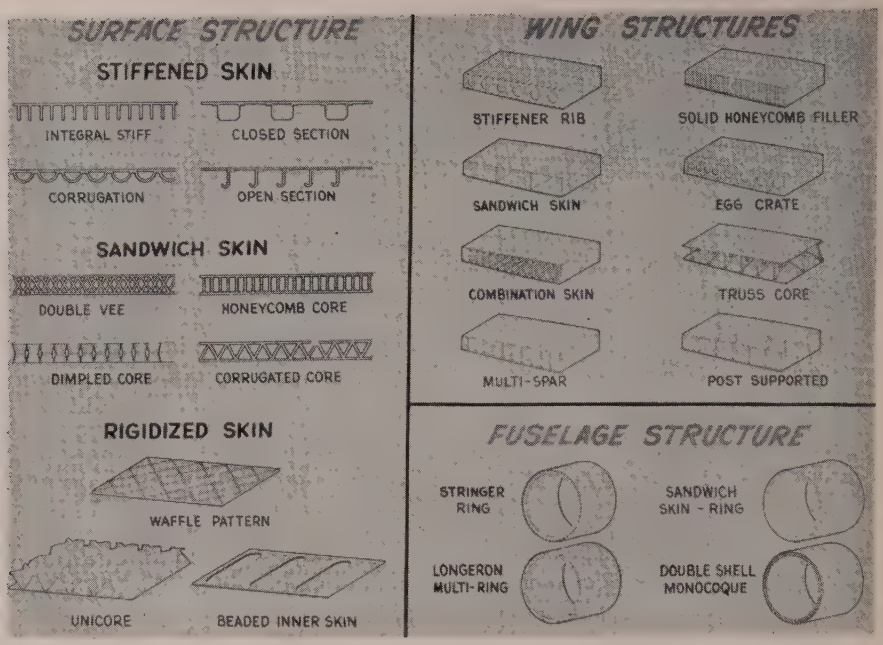


Fig. 10 — Basic structural types

ture and hence lower thermal stresses. Specific heat and thermal conductivity are related by the thermal diffusivity, high values of which are desirable. On the other hand, the stresses which result from the thermal gradients within the structure are a direct function of the modulus of elasticity multiplied by the coefficient of thermal expansion. Low values of this expression are desirable to reduce these effects.

The numerical values typical of various types of materials are shown in Table 5. The values actually vary with temperature so these shown are for an average condition. Although the diffusivity for the stainless steel is more favorable than for the titanium, the stresses are relatively insensitive to this difference and the difference between the $E\alpha_m$ values are of much greater significance. From these figures it is apparent that the thermal stresses expected for titanium is about one-half those for stainless steel. The investigation made by Hubka[10] confirms this conclusion.

Miscellaneous Properties — Finally, cost and availability of the material must be considered. The amount of money which can be spent on the material is a complex function of the other properties listed in Table 5. It is important, therefore, that the material selection is not too strongly influenced by the "price list" cost. It is obvious that the material must be available but caution must be exercised to be sure the material is *really* available in production quantities of reasonable size sheets and with stabilized and reliable characteristics. There have been cases of production delays which resulted from the necessity for ironing out material production problems which came to light during the production run.

Structural Configuration

In the past, a variety of structural configurations have been proposed and tried out. Some of these have been successful and became in a sense "standardized" in that many designers used similar types of structure for their airplane. Typical of such industry wide acceptance is the use of skin-stringer-ring type of structure for fuselages and stiffened skin-rib-beam structure for wings. Other configurations are being used where specialized conditions make these configurations suitable. The emergence of honeycomb core sandwich structure, multibeam-plate wing structures, and ring stiffened monocoque fuselages are typical examples.

In the design of a transport to operate in the Mach 3 range, it is well to review again the numerous possibilities open to the designer, even those which did not prove to be popular when applied to the slower aluminum airplane. Fig. 10 illustrates most of the basic types of structural configurations to be considered.

In considering which type of structure to choose for a specific application, it has been made plain in the previous discussion that there are many factors which play either an important or minor role. It is not possible in the limited scope of this paper to discuss in detail all these factors so a few which contribute heavily to the problem will be selected.

Effect of Density — The material property which is probably more important than any other in establishing the structural configuration is the density. All the materials which are suitable for use in this temperature range, except beryllium, are denser than aluminum alloy. This means that in order to keep the structural weight within the economic weight bounds, the material must be used in thinner gages and with smaller cross-section areas.

When used in tension, this poses the problem of having room enough for bolts, rivets, and other attachments, especially since the bolts required are the same strength as those which have been used in the past for dural structures and hence, are the same size. The thinner gages will also pose problems at joints and seams.

The major effect of the high-density materials is in that portion of the structure which carries compression or shear. In the normal case this may constitute 50–75% of the total structure. In order to maintain maximum structural efficiency and to

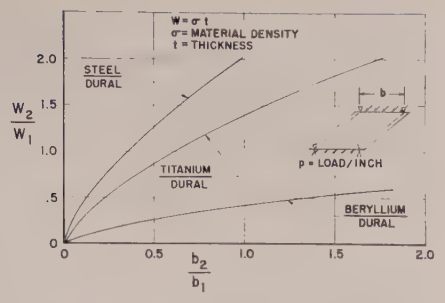


Fig. 11 — Relative skin support-weight parameter

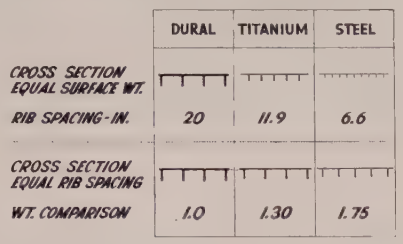


Fig. 12 — Comparison integrally stiffened surface

provide a smooth aerodynamic surface, it is necessary to prevent the buckling of the skin which, for the thinner gages, requires close support of the skin stiffening elements. This is apparent from the classical formula:

$$F_{cr} = KE(t/b)^2 \tag{11}$$

where:

F_{cr} = Buckling stress, psi
K = Fixity factor
E = Modulus of elasticity, psi
t = Skin thickness, in.
b = Distance between supports, in.

To support the same load per inch of material for two different materials:

$$(F_{cr})_1 \times t_1 = (F_{cr})_2 \times t_2 \tag{12}$$

The weight of the skin will be:

$$W_1 = t_1 \times \sigma_1 (\text{psi})$$
$$W_2 = t_2 \times \sigma_2 (\text{psi}) \tag{13}$$

σ = Density of the material (lb/cu in.)

From Eqs. 12 and 13:

$$\frac{(F_{cr})_1}{(F_{cr})_2} = \frac{W_2/\sigma_2}{W_1/\sigma_1} \tag{14}$$

From Eqs. 11 and 14 it can be shown that:

$$\frac{b_1}{b_2} = \sqrt{\frac{E_1}{E_2}} \sqrt{\left(\frac{W_1\sigma_2}{W_2\sigma_1}\right)^3} \tag{15}$$

This equation plots as shown on Fig. 11 using dural as a basis for comparison. It is apparent from this curve that when support structure spacing is being considered, titanium is superior to the stainless steel. This curve also shows one reason why beryllium has such possibilities since the support structure can be spaced very far apart and the number of parts is correspondingly reduced.

A pictorial representation of the effect of density is shown on Fig. 12 which gives the relative sizes of an integrally stiffened surface skin when made of dural, titanium, and stainless steel. Indicated also is the effect on the internal supporting structure.

Much time and thought are put into the preliminary design phase of the development to devise ways of reducing both the complexity and the weight. One approach which is very attractive is to make the surfaces of some sort of sandwich. This is equivalent to an increase of the effective thickness of the surface structure without a corresponding increase in weight. This then allows wider spacing of internal structure and the number of parts is correspondingly reduced. This approach has much in its favor and is being used extensively in certain aircraft, both supersonic and subsonic. At present, however, sandwich construction is expensive, particularly in the materials being considered for the supersonic transport — which makes it worthwhile to investigate other approaches, such as the conventional stiffened skin type of structure.

The following sections will discuss in a general way some of the factors which are useful in evaluating the suitability of the two approaches for specific application to a commercial transport.

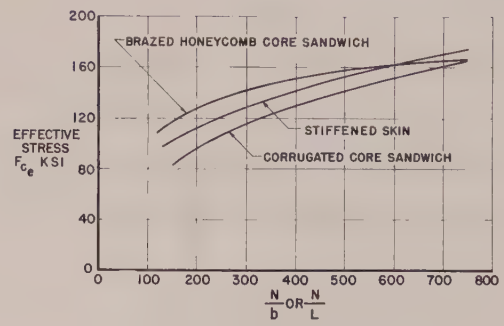


Fig. 13 — Effective allowable compressive stress (material PH 15-7 Mo, RH 950)

Table 5 — Thermal Properties of Materials

Material Type	Specific Heat C_p	Thermal Conductivity k	Thermal Expansion a_m	Density σ	Modulus of Elasticity E	Diffusivity a	Thermal Stress $a_m E$
Stainless Steel 17-7PH	0.13	0.24	6.1	0.276	27.0	0.67	165
Titanium 6A1-4V	0.14	0.12	4.8	0.163	15.0	0.53	72
Nickel Inconel X	0.115	0.19	7.6	0.300	29.5	0.55	225
Beryllium AMV-Hot Pressed	0.55	2.06	7.5	0.067	42.5	5.60	320
Aluminum 2024-T4	0.22	1.82	12.5	0.100	10.0	8.30	125

Wing Structure

The types of wing structures to be compared are:

1. Skin-stringer multirib.
2. Brazed honeycomb core sandwich multibeam.
3. Corrugated core sandwich multibeam.

These will be examined to determine in a general way how they compare on the basis of the structural design parameters given in Table 1.

Surface Structure — Compression Allowables — The allowable stress is a function of the type of surface structure, rib or beam spacing, and the temperature of the material. (Note that the effects of temperature gradients are considered to be a part

of the stress pattern in the structure). One way of comparing the allowable stresses for the different types of surface structure is to plot the allowable effective stress against the load per inch divided either by the beam spacing b, applicable to sandwich surfaces, or by the rib spacing L, applicable to stiffened skin surfaces. Such a plot is shown in Fig. 13. Room temperature properties are assumed since, in this comparative study, the decrease in allowable will be the same proportion for each type of structure.

From these theoretical curves the following observations can be made:

1. It is possible to design a stiffened skin surface structure which is competitive in weight with the brazed honeycomb sandwich structure. For instance, for a surface load intensity N of 6000 psi, a rib spacing of 15 in. supporting a stiffened skin structure results in the same allowable effective stress as a brazed honeycomb sandwich panel with a beam spacing of 20 in. Lower load intensities result in relatively closer rib spacing than beam spacing, but for the higher load intensities the rib spacing approaches the beam spacing.

2. The corrugated core sandwich structure is not as efficient as either the honeycomb core or stiffened skin structure.

No tests have been made on the specific configurations and materials shown on this curve; however, tests were made on a series of panels as shown in Fig. 14. These panels were "handbuilt" and were not particularly smooth, and some of the spot welds were substandard which resulted in local failures prior to ultimate load. In spite of these deficiencies the results of the tests encouraged us to believe that theoretical values of allowables could be achieved. Panel B displayed an interesting characteristic in that local buckles did not precipitate catastrophic failure of the panel. Panel D was an attempt to combine some of the desirable features of a sandwich panel and a stiffened skin. The results of the tests on this configuration were disappointing and for the present at least it has been dropped as a possibility.

At 600 F test temperature both Panel B and Panel D were materially affected by the thermal stresses resulting from the temperature gradient between the skin and the stiffeners. This effect was particularly severe on Panel D. These affects are calculable, however, and can be considered in the design of the structure.

Rigidity — One of the serious problems which confronts the designer is the rigidity of the structure. As mentioned previously, the flat shapes will tend to be flexible and this in turn results in numerous dynamic and loads problems. The flexibility and dynamic characteristics are affected by thickness ratio, airfoil shape, sweep, powerplant location, and other configuration possibilities. The aerodynamics of the airplane is also affected by these same parameters, and often are in the opposite favorable sense. Hence, a study must be made to evaluate the "trade off" between weight required for rigidity, weight required to obtain satisfactory dynamic characteristics other ways, and the effect on aerodynamic efficiency.

Fig. 15 shows the result of one such study on the effect of thickness ratio and aspect ratio on the weight required for adequate rigidity over that

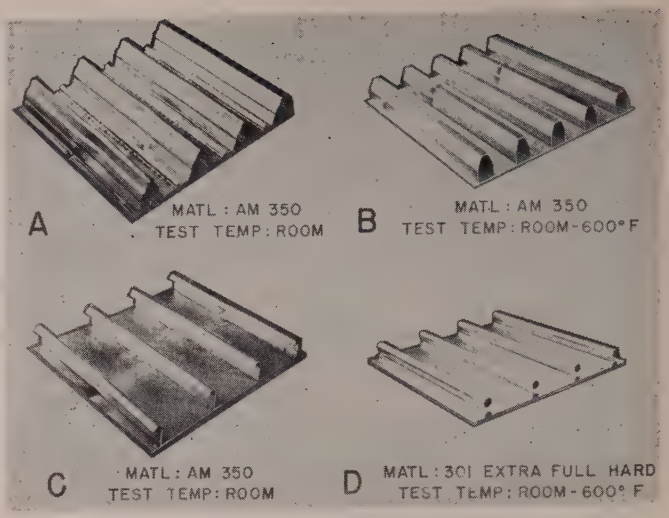


Fig. 14 — Test panels

necessary for strength alone. This curve indicates a large increment of weight can be expended to achieve rigidity. The story is not complete, however, with just this information since it must be related to the overall system. Fig. 16 is the result of a study which shows the effect on direct operating cost of thickness ratio and aspect ratio changes taking into account the weight and aerodynamics.

In addition to the external changes which will affect the rigidity, there are certain optimum ways of distributing a given amount of structural material to achieve the greatest rigidity. The best way of studying these variations is by means of analog or digital computers. However, to illustrate one aspect of the problem a simplified study was made to compare the rigidity of various types of surface structure. The types compared are:

1. Skin — stringer with skin area 70% total area.
2. Honeycomb core sandwich.
3. Corrugated core sandwich.

The combined effects of surface shear rigidity and the bending rigidity are included in Fig. 17 which is the result of this study. Including the bending and shear rigidity together is valid for a wing of this type since differential bending accounts for approximately 50% of the total torsional stiffness, From this comparison it is seen that it is possible to achieve approximately the same rigidity to weight ratio, no matter which type of structure is used.

It should be mentioned that when rigidity is the criterion, titanium alloys offer no advantage over stainless steel and in fact may be a disadvantage since the modulus of elasticity-density ratio is less favorable for titanium.

Internal Structure — The weight and complexity of the internal structure required to support the surface and transmit spanwise and chordwise shears may influence the decision of which type of structure to use. Consequently, it is necessary to look at this aspect of the structural configuration and fit it into the overall picture. Essentially two types of internal structure have been considered. The first of these is the multibeam type and its variations which goes with the sandwich surface structure, and the second is the multibeam type which

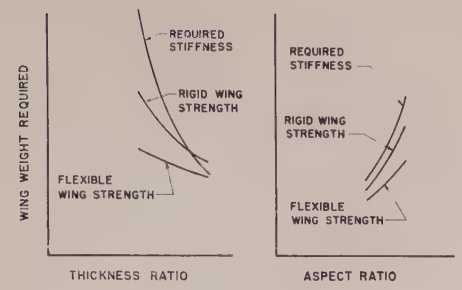


Fig. 15A — Effect of thickness ratio on wing weight (constant aspect ratio)

Fig. 15B — Effect of aspect ratio on wing weight (constant wing thickness)

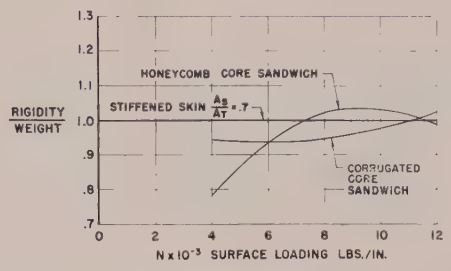


Fig. 17 — Relative rigidity-weight ratio for surface structure compared with stiffened skin with $A_s/A_t = 0.7$.

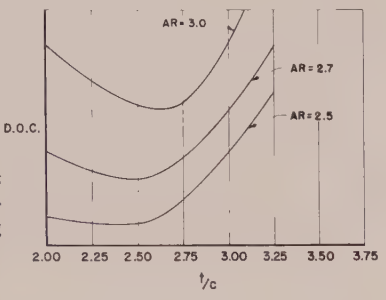


Fig. 16 — Effect of thickness and aspect ratio on direct operating cost

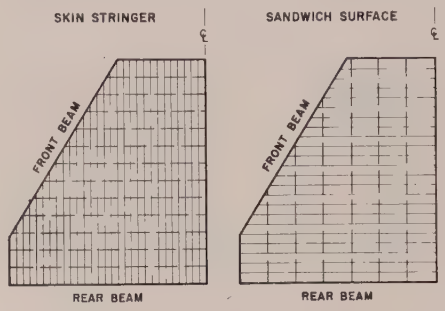


Fig. 18 — Schematic of internal structure

goes with the stringer-skin surface. Actually due to the low aspect ratio of the wing, the multibeam type requires a certain amount of chordwise support to provide shear stiffness in this direction, and the multirib type requires a number of spanwise beams to moderate the shear lag effects. A plan view of the internal structure for the two cases will look something like the schematic shown in Fig. 18. Although the diagrams look very similar, the multibeam structure has fewer parts and hence may tend to be less complex. From a weight standpoint, a preliminary analysis revealed no significant difference when account is taken of the weight required to attach the honeycomb surface to the sub-structure and to itself. Access to the interior of the wing for purposes of fabrication and inspection must also be considered and the necessary joints in the stiffened skin structure are relatively simple and easy to design to achieve minimum weight, particularly since there has been considerable experience with this type of construction.

Leading Edge and Trailing Edge — Only brief mention will be made of the leading edge and trailing edge, although the razor-thin structure presents a serious rigidity and space problem. Here the use of honeycomb cores between the top and bottom surface appears to be the best way of meeting the conditions. However, as the wing gets thicker, such a surface-to-surface core becomes heavy and other ways of building the structure become attractive. Such possibilities include regular honeycomb core sandwich, corrugated core sandwich with the core running perpendicular to the leading or trailing edge, and skin-corrugation or skin stiffener. It is not possible to state categorically which is best since the environmental conditions, that is, loads, sonic pressure levels, temperature gradients, and the like, strongly influence the choice.

Miscellaneous Considerations — Space does not permit a discussion of other problems which must be considered. Among these are thermal stress alleviation which requires internal structure such as beams and ribs to be flexible to thermal deflections yet able to carry the structural loads; surface smoothness which is achieved by close spacing of stiffeners or by suitable cores in sandwiches; thermal deflections which can cause sizeable reductions in allowables and rigidity; thermal conductivity into the wing interior, particularly if cruise fuel is carried in the wing; the effect of long life requirements, that is, tens of thousands of hours, on such design considerations as fusion welding, material stability, and joint design.

Fuselage Structure

The design conditions for the fuselage are considerably different than those encountered for the wing. Since the fuselage is essentially the container for the payload, that is, the passengers, and since the passengers are notoriously fussy about their environment, the type of structure must take account of this fact.

In early studies many ideas were suggested for fuselage structures. These ideas ranged from shielded and cooled dural structures, thick sandwich structures where the core was a structural insulating material, double wall shells, hot structures, and others. These ideas were investigated and subjected to critical appraisal relative to such factors as: Time of flight which affected the amount of insulation required for the purely insulated structure; cooling system reliability, weight, and maintenance for cooled structures; differential expansion for double wall constructions; and development time for unusual types. As the result of this appraisal, the decision was made to pursue the conventional type structures operating at the appropriate wall temperature. The other possibilities are not being com-

pletely ignored, however, and the research program at Lockheed includes investigation of these and other types.

For the particular configuration being considered, the fuselage is strength-critical, so serious consideration is given to the material to be used—titanium or stainless steel. As indicated in Fig. 11, the use of titanium results in a considerable saving in weight but there is an increase in cost. This cost penalty is well within the allowable value of a pound so that for the present the use of titanium is contemplated for the fuselage structure. This means that the structure will be a conventional stiffened skin-ring type of structure with both the stiffeners and rings spaced closely enough to prevent skin buckling. Brazed titanium sandwich structures have not been developed to a satisfactory level of reliability or cost to be considered at present.

Detail consideration must be given to ring depth which entails an analysis for fuselage shell stability. In addition to the normal problem of minimizing the ring weight, the ring is a thermal short circuit through the insulation; and the deeper it is, the more difficult it becomes to prevent thermal leakage into the cabin cooling system. Since the design internal pressure is high, the structural designer would like to use this pressure to help stabilize the fuselage shell and thus reduce the ring size, but most of the shell is designed by ground conditions or low altitude flight where the pressure is either zero or is very low and hence of not much assistance.

Another major problem which confronts the structural designer is providing a satisfactory level of fatigue cracking resistance and the prevention of crack growth beyond safe limits. One way of dealing with this problem is to restrict the tension stresses in the fuselage which results from the pressure. This can, of course, result in excessive weight since material gages are available only in discrete increments unless the more expensive tapered sheets are used. Other ways include the use of straps or bands at intervals to act as crack retarders, or frequent skin joints which tend to stop the crack growth. The more closely spaced stiffeners and rings required for the structure will also provide a beneficial effect. Considerable test work is required on specific configurations in order to be assured of a satisfactory solution to this problem.

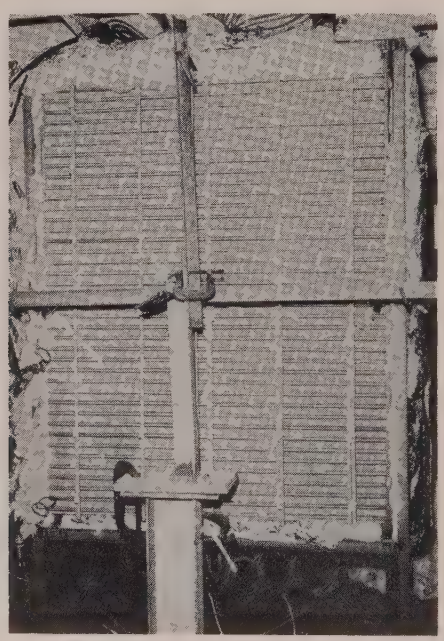


Fig. 19—Fuselage wall thermal test panel

Thermal expansion and its effect on floor supports, interior trim and cooling panels, and wing-fuselage intersection, and the transmission of heat to the interior through the structural joints are typical of problems which exert a strong influence on the detail structural design. Fig. 19 shows a test being made on a panel to evaluate, among other things, the thermal deflection and temperature distribution of a portion of the fuselage shell structure.

Other Structure

The discussion in this paper has concentrated on the wing and fuselage because the problems connected with these structures are for the most part common to the rest of the airplane. There is, however, another major structural item which presents a few unique problems and that is the engine inlet ducts which are long, hot, thin, and under high pressure. In addition to these difficult design conditions are added the complications of large movable parts and bleed openings. In short, the inlet ducts present a challenge to the structures designer which should keep him interested for some time.

Cost

The previous discussion has touched on allowable cost per pound of weight saved, the weight ratio, materials, structural configuration, and wing and fuselage structures. The time has now come to discuss briefly the effect of these on the airplane and its cost.

A considerable amount of time and effort has been put into studying the advantages and disadvantages of both brazed honeycomb sandwich structures and the stiffened skin structures. Among the reasons for the emphasis which has been placed on the more conventional stiffened skin types is that if called upon to produce an airplane today, this type of construction would lend itself more readily to in-plant fabrication with a minimum amount of increased capital equipment. Also the development time would be less and, in view of the present-day high cost of sandwich construction, the airplane could be produced more cheaply. For example, current estimates of cost indicate a factor of 3 or more on the basic structure if made of brazed honeycomb sandwich. This would amount to a few million dollars on the initial price, which makes it worth spending considerable effort in evaluating the alternates. This cost factor is, of course, time-dependent and all the effort being put into the development of brazed honeycomb sandwich is bound to bear fruit and the future will see a substantial decrease in this cost.

In view of this probable reduction in cost, serious consideration must be given to the possibility that the two types may become competitive costwise. In this event, would the stiffened skin structure still hold a competitive position as an efficient structure? Studies to date indicate that in the overall system of the type of airplane and mission being considered it is competitive. Other types and other missions could well result in a different conclusion. There are, of course, a considerable number of places where sandwich structure will be used in any event, places

where nothing else will do the job or where the advantages pay for the cost, and as the price comes down, the number of applications will increase. As time goes on and the design evolves, there must be a constant re-evaluation of the current facts and future possibilities to insure a final product which incorporates the latest technical advances.

It was previously mentioned that the use of titanium in place of stainless steel also results in increased cost. In this case there is a sizeable weight pay off and in certain portions of the structure, where the additional cost has been estimated to range between $75 and $100 per pound of weight saved, the use of titanium is justified since it is well within the maximum allowable of $350 previously determined.

Service Utilization

The last topic to be discussed very briefly is the effect of structure on the utilization; a topic of great interest to the operator. It must be kept in mind that the structure chosen must be inspected, maintained, and occasionally repaired. The airline which will perform this task is faced with a problem similar to that encountered by the manufacturer — a new material more difficult to handle and which requires somewhat different machines and facilities than are at present on hand. Although it is likely that techniques in the inspection and repair of honeycomb sandwich will be developed, it appears that the open type of structure, as exemplified by the stiffened skin approach, offers the operators a lesser problem from an inspection and maintenance standpoint. This consideration then presents another point in favor of the more conventional structural approach.

Summary

A general discussion of some of the problems facing the structures engineer as he considers the design of a Mach 3 transport has been presented. These problems relate generally to the aerodynamic heating, high altitude, and thin shapes inherent with an airplane of this type. Definitive solutions to these problems have not been attempted. Indeed it is not desirable to do so at the present time for technological advances are being made so rapidly that maximum flexibility must be maintained as long as possible.

The day will come, however, when a decision must be made and the design frozen. Some important design parameters have been presented which must be compromised in as near optimum fashion as possible. These relate to the structural efficiency, initial cost, and service utilization.

The value of a pound of weight and the meaning of the weight ratio has been discussed and approximate values suggested for purposes of guidance. It is important to realize that for a supersonic transport, the value of a pound is increased by a large factor over those previously considered for the subsonic airplanes.

Material properties and their effect on the structure has been discussed and various materials have been compared as to their suitability. It can be concluded that no one material presents the perfect solution but that a variety of materials will be required.

The structural configuration of the wing and fuselage is discussed. It is pointed out that the wing may be rigidity-critical rather than strength-critical in which case the configuration will aim toward the maximum rigidity-weight ratio. The fuselage on the other hand may be strength-critical and for this case stiffened skin construction using titanium material becomes very attractive.

Finally, a short discussion of cost and maintenance is presented. The cost difference between a conventional stiffened skin and brazed honeycomb sandwich at the present time is large and gives the stiffened skin construction an advantage. Maintenance considerations also favor the use of conventional structure.

Additional References

Materials —

1. "Performance of Materials at Elevated Temperatures for Aircraft and Missile Applications," by Alan V. Levy. Marquardt Aircraft Co.
2. "Comparison of Primary Structural Materials for Use at Elevated Temperatures," by J. E. Rhodes. Report No. 10464, Lockheed Aircraft Corp.
3. "Trends in High-Temperature Metals," by Julius J. Harwood. *Aircraft and Missiles Manufacturing,* May and June, 1958.
4. "Stainless or Titanium for Supersonic Aircraft," by R. G. Sloan. *Western Aviation,* October, 1956.
5. "Workhorses of the Industry, PH Stainless Steels," by J. N. Barnett and W. W. Rhode. *Aviation Age,* R & D Handbook.
6. "Status of High-Strength Steels for Aircraft Industry," by R. J. Nekervis, C. H. Lund, and A. M. Hall. Battelle Memorial Institute TML Report No. 91, Jan. 3, 1958.
7. "Material Properties for Design of Airframe Structures to operate at High Temperatures," by L. R. Jackson. Titanium Metallurgical Laboratory, TML Report No. 38, March 23, 1950.
8. "Crack Propagation Tests on Several High Strength Steels," by T. J. Stapleton. L. R. 13207, Lockheed Aircraft Corp., California Division, Feb. 2, 1959.
9. "Strength of Metal Aircraft Elements," MIL-HDBK-5 Coordination Copy.

Allowable Stresses —

10. "Some Research Results on Sandwich Structures," by Melvin S. Anderson and Richard G. Updegraff. NACA TN 4009, June, 1957.
11. "Local Instability of Elements of Truss-Core Sandwich Plate," by Melvin S. Anderson. NACA TN 4292, July, 1958.
12. "Engineering Status Report of Budd Integrated Core Panels," Vols. 1 and 2. U. S. Navy Bureau of Aeronautics Contract NO as58–622C. The Budd Co.
13. "Buckling of Flat Sandwich Panels under Edgewise Compression or Shear Loading." Report No. ER-3282, Lockheed Aircraft Corp., Georgia Division.
14. "Local Instability of Sandwich Construction under Various Types of Loading." SMN 37, Lockheed Aircraft Corp., Georgia Division.
15. "Stress Memo Manual." Lockheed Aircraft Corp., California Division.
16. "Theory of Elastic Stability," by S. Timoshenko. Pub by McGraw-Hill Book Co., Inc., New York, 1951.

The Design for Better Ride

H. O. Flynn
Chevrolet Motor Division, General Motors Corp.

THE PURPOSE of this paper is to describe the aspects of an appropriate chassis design which will provide improved handling, ride, and driver comfort, with increased load capacity.

The basic theory of such a design has been covered in a previous paper.[1] Because of "product security" it was limited to some of the theories of handling. With the new product on the highway, all phases of ride, handling and their corollary, driver comfort, can be discussed freely.

Vehicle suspensions, in eliminating the more damaging effects of direct shock, introduce an entirely different set of motions which involve frequencies and amplitudes. Any study of comfort naturally implies that a correlation exists between these motions and the degree of human tolerance.

Since 1931 there has been an ever-increasing volume of research conducted to gain a better understanding of the effects of vehicle vibration upon man. Most of this work has been directly related to an objective evaluation of passenger discomfort and tolerance to vibration in various systems, from vibrating platforms to supersonic test sleds. Most of the data, as applied to moving vehicles, is of no real value because the subjects were standing or lying prone on tables, while those seated were usually positioned on what may be termed a board. The conventional approach has been to submit subjects to sinusoidal oscillations at various amplitudes and then vary the frequency. Evidence and data were compiled from the subjects' expression of like or dislike of the motions experienced.

In assessing the overall effect of vibratory motion on human beings, it must be remembered that the effects are both physiological and psychological. In analyzing data gathered over the years it was necessary to interpret the pertinent points from an engineering standpoint, since either physiological or psychological interpretations require special training not common to an engineering department.

Shown in Fig. 1 is a "Meister thresholds" curve, plotted in Berlin in 1935. The results were obtained

---

A PRIMARY REQUIREMENT of today's trucks is that they carry loads quickly, dependably, and profitably. Until recently, a high load capacity and a reasonably comfortable ride were incompatible — load requirements called for suspensions adversely affecting driver comfort. This paper describes one manufacturer's approach to the problem.

The author has divided his paper into several sections. The first is a general discussion on the nebulous concept of comfort. He then describes the broad engineering approach required to attain increased comfort for personnel and cargo. The third section contains a detailed description of the specific Chevrolet approach to obtain the goals of increased comfort. The final two sections deal with testing and the special instrumentation used to evaluate the new truck models.*

# of a Truck and Handling

with the subjects standing and exposed to vertical vibration. Data is plotted to represent a range between the strongly noticeable and extremely uncomfortable. This report is of particular interest since it adequately covers the important low frequency range below 8–10 cps.

It is evident that an increasing dislike was shown for greater amplitudes at a constant frequency and vice versa.

Fig. 2 is a curve of comfort, which was compiled by Maurice Olley at the GM Proving Grounds in 1934. Subjects were tested on a bouncing table, having a cushioned seat and foot rest. Again, the results confirm other findings in that both high frequencies and high amplitudes produce discomfort.

Fig. 3 is the "Jacklin limits" curve compiled by Jacklin and Liddell, at Purdue University in 1933. In testing the subjects, the amplitude setting was fixed and the frequency increased until it became disturbing or uncomfortable. The tests covered a frequency range of about 1–7 cps, with amplitudes up to 1¼ in. This investigation is one of the most extensive available.

Still other investigations are shown on this composite curve (Fig. 4). It is interesting that all studies generally fall within the thresholds of comfort compiled by Meister in 1935.

These are but a few of many hundreds of studies conducted over the years in many sections of the world. The most striking aspect is the close correlation of all conclusions that tolerance to greater amplitude necessitates decreased frequency and, conversely, that higher frequencies must be accompanied by smaller amplitudes.

It is common knowledge that the ride motions of today's trucks fall, to a large degree, into the dis-

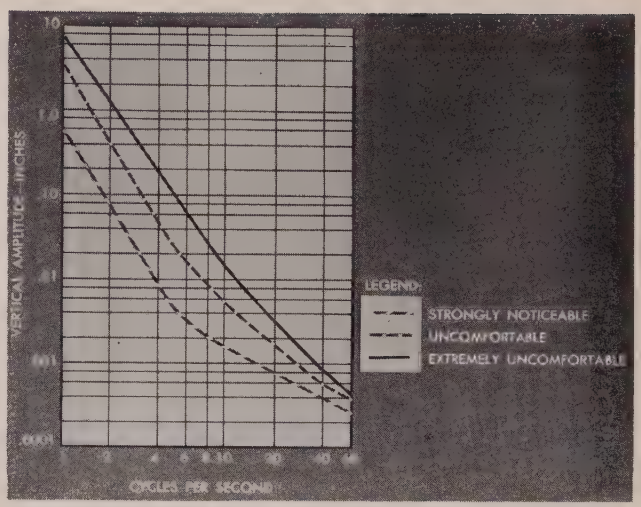


Fig. 1 — Meister thresholds, vibration amplitude versus frequency for constant sensations (F. J. Meister, May-June, 1935, Forschung, U.D.I.Berlin, Germany)

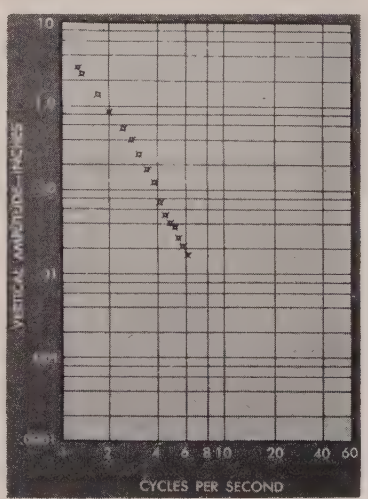


Fig. 2 — Curve of comfort, vibration amplitude versus frequency for constant sensations (M. Olley, March, 1934, General Motors Proving Grounds)

* Paper presented at SAE National Transportation Meeting, Chicago, Oct. 28, 1959.

[1] "Road Behavior of Modern Truck," by A. C. Mair. Paper presented at SAE Summer Meeting, Atlantic City, June, 1957.

comfort zone. Three solutions are possible. One is to reduce frequency while maintaining amplitude, another to reduce amplitude while maintaining frequency, and the third to reduce both.

Truck design and periodic improvements have, over the years, been primarily concerned with increased vehicle capacity and durability. An expanding economy and increased demand for greater production of vehicles dictated that many things be postponed until a more expedient time. It was inevitable that a compromise be made between the aspects of ride, handling, and other factors.

There are so many variables that enter into truck design that until just recently, a high load capacity and a reasonably comfortable ride was incompatible. High load capacity requirements led to designing suspensions which adversely affected driver comfort; improvements on one could be made only at the expense of the other.

The primary requirement of today's truck is that it carry loads quickly, dependably, and profitably, within the legal restrictions. Until now truck designers have incorporated into their products what consumer demand required. It was inevitable, however, that eventually both ride and handling improvements would have to be considered along with durability and payload potential.

Ride

The program was to incorporate several factors for further study and evalution.

A prime consideration was to provide for heavier payloads at higher safe road speeds with new and improved ride, handling, and driver comfort characteristics. These would reduce truck and payload damage, with increased efficiency to the consumer and, in turn, increased consumer acceptance.

After preliminary investigation, it was obvious that new and radical changes would have to be made to the generally accepted suspension design of today's trucks, since the I-beam and high rate springs had reached the optimum in the compromise of ride and carrying capacity.

Ride of a vehicle has two basic aspects: (1) a high frequency, low amplitude motion induced by wheels moving over irregular road surfaces; and (2) the lower frequency, but higher amplitude bounce and pitch motions of the suspended mass vibrating on its spring system.

Bounce motion occurs when both ends of the vehicle move in the same direction about an oscillation center outside the wheelbase. Pitch motion occurs when each end of the vehicle is moving in an opposite direction about an oscillation center within the wheelbase.

Items which produce the greatest effect on ride are the deflection of the spring system (springs and tires), the inertia effect of the sprung mass, and the wheelbase. These things determine the bounce and pitch frequencies, and their respective oscillation centers, or what can be termed the overall character of the ride.

Probably the most important is the spring system deflection (Fig. 5). This concerns the total amount of deflection of the tires and chassis springs that occurs between the unsupported position of the sprung mass, and the position assumed after the sprung mass is completely supported. It is measured at the wheels for both front and rear.

Since it applies only for a constant rate spring system which rarely exists, the effective deflection must be used, which is the sprung load divided by the wheel rate. This, in turn, produces a value equivalent to a system having a constant rate.

The factor next in importance is the flywheel ef-

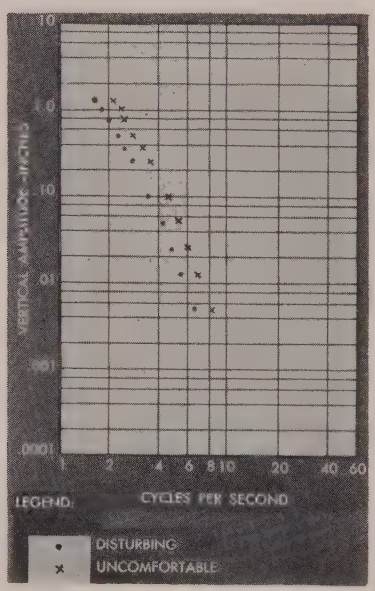


Fig. 3 — Jacklin limits, vibration amplitude versus frequency for constant sensations (H. M. Jacklin and G. J. Liddell, May, 1933, Purdue Univ.)

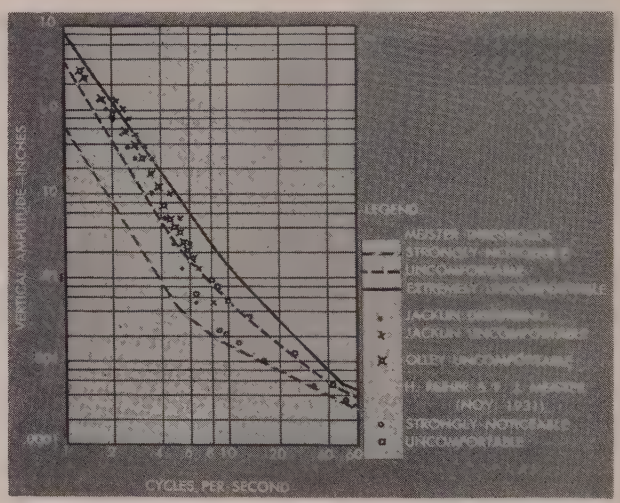


Fig. 4 — Overall correlation, vibration amplitude versus frequency for constant sensations

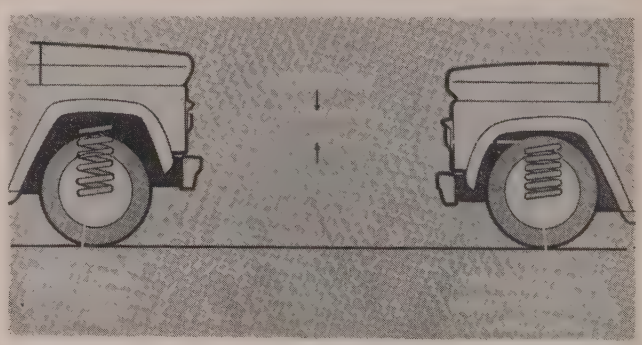


Fig. 5 — Spring system deflection

fect of the sprung mass which involves the radius of gyration, the location of the center of gravity (c.g.), and the length of wheelbase. This relation between flywheel effect and wheelbase can be written K^2/ab and is known as the K^2/ab ratio (Fig. 6). K is the radius of gyration, a is distance between c.g. and front axle and b is distance between c.g. and rear axle.

The radius of gyration of the truck is affected by the distribution of its mass, that is, if the mass is concentrated high or at the ends of the truck, K is large. Conversely, if the mass is concentrated near the c.g., K remains small. With this in mind, it is apparent that with the greatly varying loads trucks carry, the various types of bodies, and the large differences in wheelbase, the K^2/ab ratio of trucks varies greatly.

The character of truck ride can be expressed mathematically and, therefore, a study can be made of various approaches to the problem, with a somewhat predictable outcome (Fig. 7). The simple expression $f = \frac{188}{\sqrt{d}}$ can be applied to determine the bounce frequencies of a single mass on one spring or of a vehicle with equal deflection front and rear and the c.g. midway between spring centers. It applies also for the pitch frequency of a vehicle with equal deflections and having a K^2/ab ratio of 1.

The first condition occurs only on a unicyle and the latter two are to be seen only occasionally. Usually, the deflections are unequal and the K^2/ab ratio may vary from 0.5 to 1.5.

Through research, expressions covering all vehicle conditions have been determined. These solve for bounce and pitch frequencies, and for the location of the bounce and pitch centers.

Fig. 8 shows how the location of the bounce and pitch centers is altered by changes in front and rear deflection and changes in K^2/ab ratio. It is apparent that the ride character of a truck changes as the load is varied. The oscillation centers move toward the end of the vehicle with the softer springs.

In the case of a typical 2-ton truck (Table 1) with an unloaded weight of 8050 lb and a loaded gvw of 19,320 lb, with the load distributed so that K^2/ab equals 1 and a wheelbase of 10 ft, the vehicle would have a front effective deflection of about 4 in. This data thus provides sufficient known quantities to solve for ride frequency information.

The first item of note is that the loaded vehicle

Table 1 — Ride Frequency Information for Two-Ton Truck

Condition	1 Loaded 19320	2 Unloaded 8050	3 Loaded 19320	4 Unloaded (Constant rate springs) 8050	5 Unloaded (Constant deflection springs) 8050
Front Spring Deflection, in.	3	2.12	9	6.4	9
Rear Spring Deflection, in.	4	.89	6	1.33	6
K^2/ab Ratio	1.0	0.725	1.0	0.725	0.725
Pitch Frequency cpm	109.3	225	76.5	183	87.2
Bounce Frequency cpm	94	134	62.5	77.8	64.5
Pitch Center from C.G., in. forward	72	23	72	26	16.5
Bounce Center from C.G., in. back	48	100	48	87.5	137.5

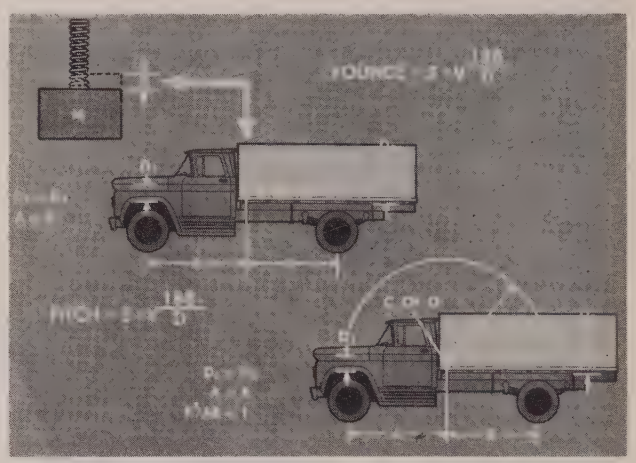


Fig. 7 — Bounce and pitch frequency

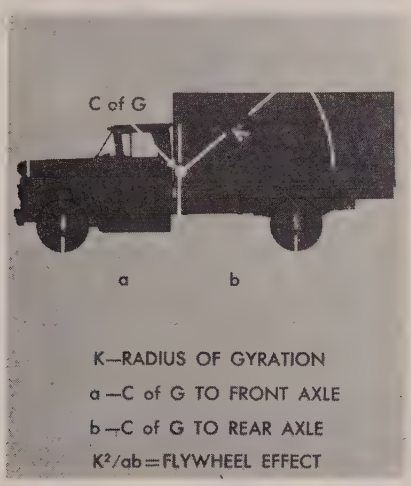


Fig. 6 — Truck flywheel effect

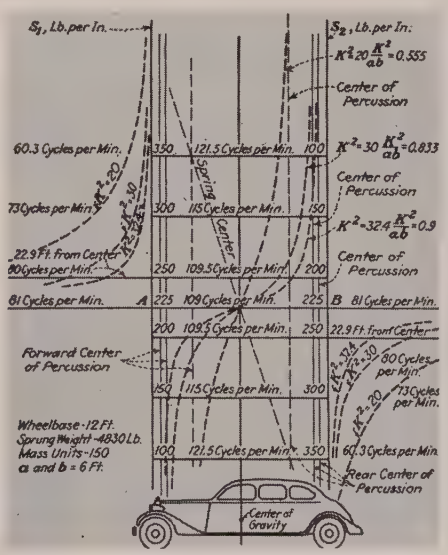


Fig. 8 — Oscillation center versus spring rate

has a pitch frequency of 109 cpm and a bounce frequency of 94 cpm. These allow very little displacement to stay within the comfort zone. The second column shows the same vehicle unloaded. The effective deflection becomes 2.12 in. front and 0.89 in. rear. This results in a pitch frequency of 225 cpm and a bounce frequency of 134 cpm.

Furthermore, the solution assumes no friction or damping, and the excesive friction existing in truck leaf springs, and in particular, the massive rear ones, is a well-known fact. The rear friction is so high that in the case of the unloaded truck with little sprung weight at the rear there is insufficient force to overcome this friction on small bumps. Thus, the vehicle bounces or pitches on its tires at a frequency which exceeds 225 cpm, and may go as high as 400 cpm.

The obvious solution is to soften the springs and reduce suspension friction. A deflection comparable to a modern-day passenger car's, which has ride frequencies below 75 cpm, would have to be above 9 in. front and 6 in. rear deflection, loaded.

Softer springs demand low suspension friction. Friction prevents spring deflection for very small wheel motions and forces the vehicle to bounce on its tires on smooth roads at extremely high and objectionable frequencies.

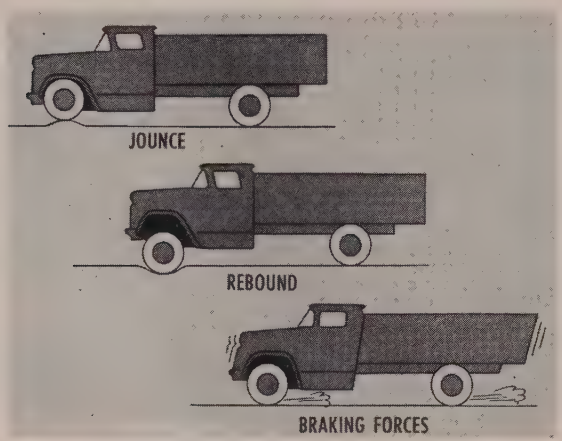


Fig. 9 — Load categories of front suspension

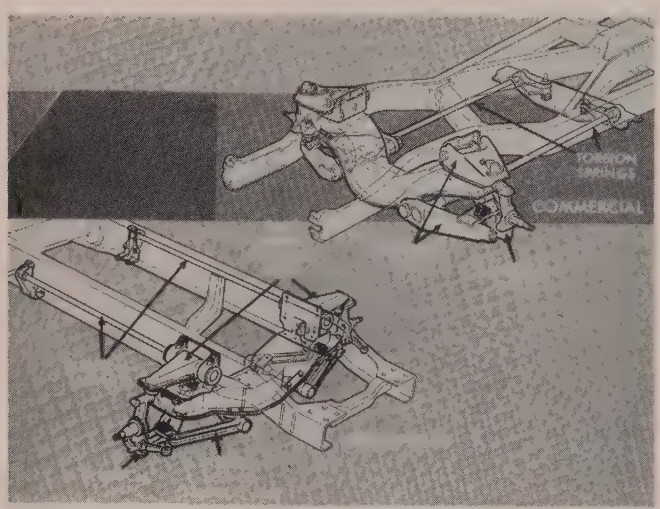


Fig. 10 — Front suspensions

The third column shows that low friction, combined with the 9 in. front and 6 in. rear effective deflection, achieves a pitch frequency of 76.5 cpm and a bounce frequency of 62.5 cpm which is an appreciable improvement.

Column 4 indicates that with this truck unloaded, the pitch frequency returns to the undesirable 183 cpm. Years of experience by all manufacturers have shown that with a conventional front axle and leaf springs, the present condition is the optimum, since the spring must serve as a locating device.

Greater deflection, like the 9 in. in the example, results in excessive brake windup and dive, difficult control of the unsprung masses, inadequate roll resistance, dangerous high-speed wheel disturbance, with the onset of the shimmy cycle.

Since the loaded truck with 9- and 6-in. effective deflection rides poorly unloaded, why not give the unloaded truck 9 in. front and 6 in. rear deflection?

With constant rate leaf springs, however, the deflection would become nearly 13 in. in the front and an unthinkable 27 in. in the rear. Wind-up, dive, and front wheel disturbance would be intolerable. The frequency might reach the seasick range. The sprung mass would be unstable, and the standing height at the rear would vary 21 in. between loaded and unloaded.

Also, steel leaf springs with this much deflection would weigh twice as much as a current production truck spring, and involve excessive use of material. An analysis of the situation shows that the ideal solution would be a spring of reasonably constant deflection and height.

The fifth column of Table 1 shows how the unloaded truck rides with constant deflection and it is evident that the problem of frequencies has been solved.

A few remarks on the higher frequency vertical and lateral shake motions are in order. These motions are less well-defined because of their complexity of sources. One factor which has a fairly stable pattern, however, is caused by the unsprung mass bouncing on the tires between the springs and the road, known as the wheel hop frequency. It is consistently in the range of 500–600 cpm. Here, springs of larger deflections are needed because forced wheel displacements result in less load increase against the sprung mass.

Thus, the obvious method of reducing frequency is to reduce the spring rate. A conventional truck chassis design makes this impossible. The front leaf springs must be stiff enough to maintain suspension alignment, as well as absorb acceleration and/or braking torques.

Design

Front Suspension — General — A thorough investigation of all aspects of softer springing indicated that Chevrolet should adopt independent front suspensions for the entire series of trucks. The linkage design would relieve the spring of all responsibility but to serve as an elastic member. Thus, springs with a much lower rate could be used and the required reduction in friction obtained.

This decision was further strengthened by the inherent benefits of an independent front suspension resulting in improved vehicle stability and handling. Independent wheel action reduces wheel wobble and

shimmy, which are basically caused by gyroscopic precessional forces acting on the wheel being displaced. Optimum roll steer characteristics are provided by positive toe-in and camber control. Further, this type of a suspension makes possible improved roll rates and dive control on braking.

Different types of independent front suspensions were considered, including Du Bonnet, long and short arm, parallel arm, trailing arm, and many others. Several methods of springing were also considered, such as transverse torsion bars, coil springs, torsion bars, and even leaf springs with nylon inserts.

It was finally decided to use long and short, parallel control arm linkage, incorporating longitudinal torsion bars as the elastic members since for many reasons this type of suspension proved most advantageous for truck usage.

Two types of independent front suspension designs are employed for the entire 1960 line. One type is used on the ½- through 1½-ton models, and another for the 2- and 2½-ton units. Both are identical in function but differ in method of construction, weight, and component location.

At the very outset of such a contemplated design, basic standards had to be established to ensure durability of components. To arrive at a baseline, a study was conducted of the heavy-duty military vehicles with independent front suspensions which saw service in the mountainous terrain of Korea; of foreign vehicles, some with independent front suspensions; and of all other independent front suspensions used in this country, including production passenger cars. This data, coupled with that already established for the I-beam suspensions, provided the necessary knowledge of safety factors, from which design could proceed.

These criteria were divided into three categories, based on the yield point of metal (Fig. 9). First, those parts which are stressed by the bump load have a safety factor of six to seven times the stress at maximum suspension capacity. Those parts which are stressed by rebound loads have a safety factor of twice the maximum suspension capacity. Those parts stressed by the braking loads have a safety factor of twice the maximum braking load that occurs on the vehicle during braking.

These considerations formed the basis for the initial stages of design. Of course, the design was not produced by these criteria alone. For all multiple load areas, it was of utmost importance to rely on the laboratory and proving ground test reports. The overall result is a homogeneous design with sufficient overall strength to provide an absolute measure of safety for even the most demanding vehicle operations.

Front suspension systems (Fig. 10) on the ½-, ¾-, 1-, and 1½-ton vehicles, which will hereafter be referred to as commercial models, are basically similar in construction. Each incorporates stamped upper and lower control arms, the same method of torsion bar attachment, and similar geometry. Ratings, however, are tailored to the vehicle size. These are 2500 lb for the ½-ton, 3000 lb for the ¾-ton, 3500 lb for the 1-ton, and 4000 lb on 1½-ton models.

The second family of front suspensions is used for the 2- and 2½-ton or heavy-duty models.

Suspensions with three different capacities are available. A 5000-lb unit is used on 2-ton vehicles, a 7000-lb suspension for the 2½-ton models, and a 9000-lb system available as an option for the 2½-ton models where severe front end loading is encountered.

The capacities are consistent with the size of the components making up the various suspensions.

Track Change — Both commercial and heavy-duty suspensions incorporate the long and short arm principle (Fig. 11) wherein the lower control arm is approximately twice the length of the other, to provide a minimum of track change during wheel movement. Theoretically, with the use of short and long control arms, it is possible to design an independent suspension with absolutely no track change as the wheel travels from jounce to rebound. This, however, holds true for only one specific rolling radius. In the case of trucks, where several tire and wheel sizes are available for each series, a compromise is necessary. The Chevrolet design provides for a maximum of ¼ in. track change during the total wheel travel, from jounce to rebound.

Camber — Another consideration governing the length of the control arm is the amount of camber change occurring during vertical wheel movement (Fig. 12). The average change on commercial models was kept within the limits of a negative 1¼ deg at full jounce, to a positive ½ deg at curb load, to a negative 2½ deg at full rebound. These smaller changes result from the suspension being very large. The heavier models have a variation in camber angles from a negative ¾ deg at full jounce, to a positive 1 deg at design height, to a negative 1¾ deg at full rebound.

Caster — Caster angles (Fig. 13), on the other hand, were determined through development. The initial commercial experimental vehicles carried

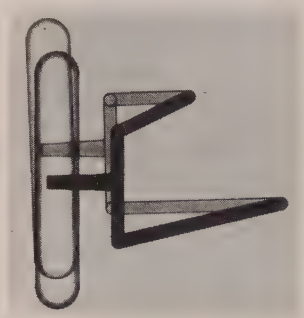


Fig. 11 — Constant tread feature with parallel-link type of suspension

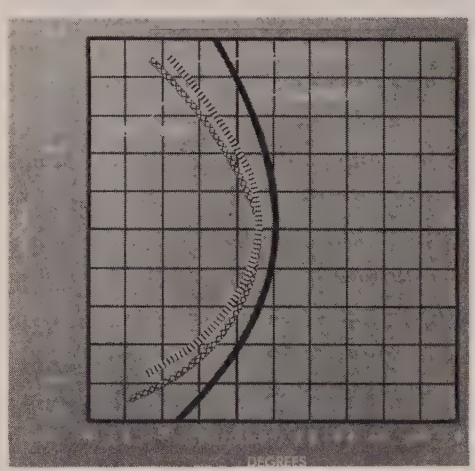


Fig. 12 — Camber angles

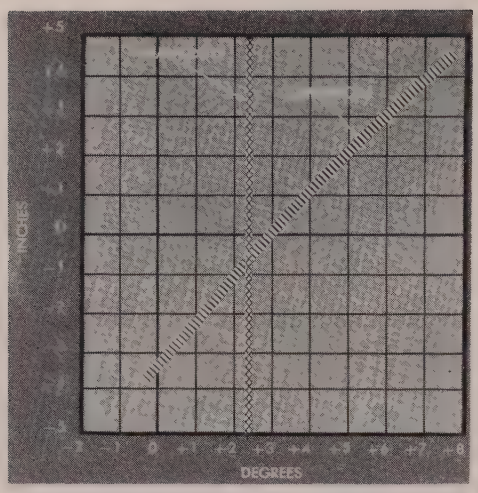


Fig. 13 — Caster angles

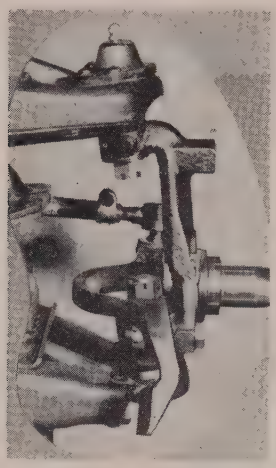


Fig. 14 — Spindle support attachment

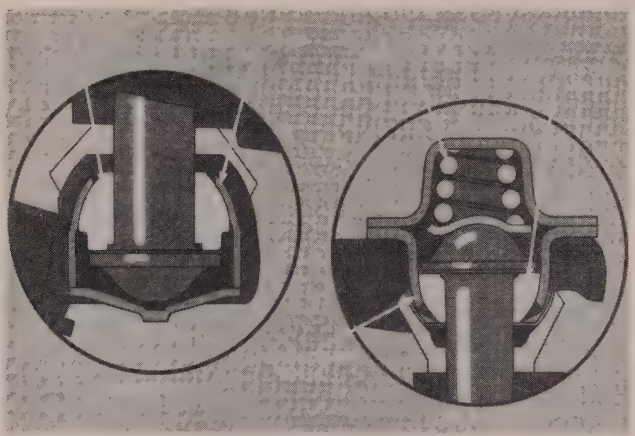


Fig. 15 — Spherical joints

Fig. 16 — Front suspension cross-members

½ deg of caster at design load, the exact duplicate of a Chevrolet passenger car with a three-passenger load. It was found that at curb weight, the rear of the vehicles raised sufficiently to decrease the caster angle by as much as 3 deg. Thus, the unloaded truck had negative caster angles of 2½ deg with resultant poor wheel return characteristics.

The caster angle range on all models was finalized at about ½ deg at curb load and about 3½ deg at design. Commercial models have a greater range than the heavy-duty units since the antidive feature produces a caster change which is due to geometry.

Spherical Joints (commercial) — Early in the design program spherical pivots were chosen for the bearings of the steering knuckles, to utilize to the utmost the advantages of independent front suspensions (Fig. 14). These were to be used on both the commercial and heavy-duty models. It became apparent that with general usage several different design conditions would be required. Thus, in the commercial unit category one set is provided for ½- and ¾-ton units, and another on the 1- and 1½-ton vehicles. In the heavy-duty classifications, both the size and design of the spherical joints is related to the suspension capacity, so that different sets are required for the 5000-, 7000-, and 9000-lb suspensions.

The basic requirement for reliability of the entire suspension is reflected in the conservative design conditions which were set up in such primary structural design elements as the pivot points (Fig. 15).

The actual bearing design for the commercial models remains quite conventional and uncontroversial. The rocking motion of both the upper and lower stud is taken by a loose, sintered iron bearing in a hardened socket, while turning takes place on a cold upset stud. The contact areas are induction hardened.

One of the important benefits derived from this design feature is the elimination of wheel wobble and shimmy tendencies, basically caused by loosely fitting spindle support components. The nonload-carrying upper ball stud of commercial models incorporates an automatic wear take-up feature as a result of the spring loaded housing, maintaining a constant relationship between the stud and its seat, which compensates for any change in size due to wear. The lower spherical joint, which retains and sustains the load in one direction, provides the self-compensating feature of eliminating any looseness of fit.

In addition, these same design features permit precise control for the friction required to obtain the desired ride.

Front Suspension Cross-Members — Elimination of the relatively rigid I-beam and leaf springs, and the substitution of a linkage suspension system, creates the need for a rigid frame member to which the suspension can be attached, and which can absorb all the forces transmitted directly from the road to the moving vehicle. The suspension cross-members (Fig. 16) are massive, hat-section structures, reinforced by a bottom plate to form torsionally rigid boxes.

The design of the suspension is such that the entire suspension unit can be preassembled, with proper camber and caster before being assembled to the vehicle.

Torsion Bars — The torsion springs (Fig. 17) were chosen because of their greater adaptability to use on trucks with independent front suspensions. In comparison to leaf springs, torsion bars will store approximately four times as much energy per pound of steel.

Coil springs were entirely out of the question. To provide the necessary carrying capacity, yet fit the space available on heavy-duty models, the wire diameter would have to go up to 1¼ in., yet the spring itself would have to have as few as three active coils. In addition, the wide range of loads common to an entire line of vehicles would demand springs of many coil diameters and wire sizes. To fit this range of diameters into their seats in the front suspension cross-member would require either separate cross-members and control arms for each spring size, or else adapters.

At present, the torsion bar material is AISI-5160, manganese chrome steel, heat-treated to Rockwell C47–50. The bars are centerless ground to size, after which the hexagon ends are uset in their correct angular relationship before prestress. The bars are then heat-treated in a controlled atmosphere furnace to control the decarburization, and are straightened if necessary. Magnetic particle testing detects any faults.

Shotpeening follows, after which the bars are preset. The presetting operation consists of applying torque in the same direction as the normal load operation of the bar to about 7% above the maximum operating stress. Preset strain equals 0.022 in. per in, with permanent set strain of approximately 0.008 in. per in. Thus, prestressing produces an angular set in the bar of about one-third the total presetting angle. This serves to increase the elastic limit of the material by raising the yield stress in the direction of preset, increasing the strength and durability of the bar, and reducing sagging.

An explanation of load-deflection and stress distribution which occurs during prestressing is taken from the SAE manual on design and manufacture of torsion bar springs.

Shown on Fig. 18 is a load deflection diagram, using a theoretical bar of 2.27 in. diameter, 72.7. in. active length with a torsional rate of 6900 in.-lb per deg. Prestressing of the bar begins at point zero and shows a linear increase in torque corresponding to its rate, up to point A. There partial yielding occurs, with the torque building up more slowly for each degree of bar windup. Point B represents the maximum safe torque and windup. As the load is removed the torsion bar returns elastically from point B, and along a linear decrease to point C. The torsional rate, however, is somewhat reduced, due to reduction in shear modulus from 11.0×10^6 psi before presetting to 10.5×10^6 psi after presetting.

At point B all material in the bar is either at the yield point or below. Therefore, as long as operational torques do not exceed the windup torque at point B, point D, for an example, being maximum, the bar will operate elastically along the line BC, with at least a 20 deg greater possible windup before the yield point is reached than before prestressing.

Because of the fact that our operational windup is but a small fraction of the prestressing windup, the benefits are somewhat self-evident. As partial yielding in prestressing does not occur instantly, Chevrolet specifications require that the bar be wound up to point B at least three separate times.

The phenomenon through which the torsion bar acquires its greater capacity is depicted in Fig. 19. Curve OAB is the same as Fig. 18, but to a different scale. The ordinate and abscissa have been refined, converting torque to stress. Stress equals torque divided by section modulus. Strain, which equals windup angle, times one half bar diameter, divided by bar length, replaces the windup angle. Incidentally, stress has no significance above point A, since it is calculated by a formula which holds true only below the elastic limit. The true stress, however, can be derived from curve OAB.

Through any point F on curve AB, a vertical line FD is drawn. A line is then drawn tangent to curve AB at point F, intersecting the y axis at point G. A line, equal in length to ¼ the length of OG, is then dropped to point C. A continuation of this process provides curve ACI. Thus, the actual stress at the surface of the bar is DC when the nominal stress is

Fig. 17 — Torsion bars

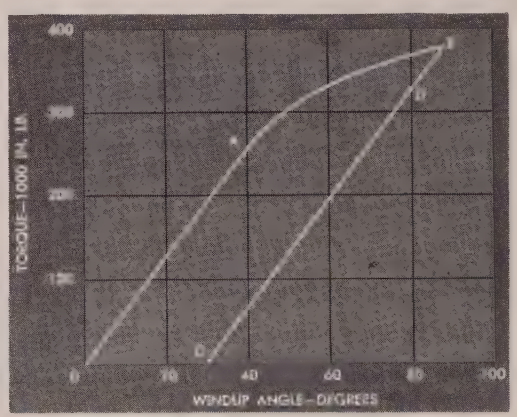


Fig. 18 — Load deflection during presetting

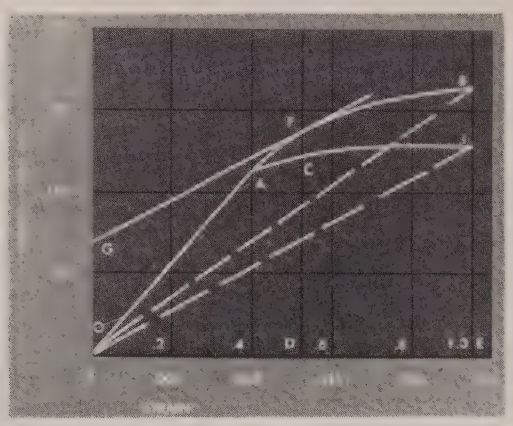


Fig. 19 — Stress-strain during presetting

DF and the strain is OD.

It is seen that the material is strain hardened from 115,000 psi at point A to 130,000 at point I, which partially explains the increased capacity.

Even more important, however, is the changed stress distribution. Considering the bar to be a homogeneous structure, any element below the surface must have the same stress-strain curve as any single particle at the surface. It is generally accepted that during the twisting of a round bar the cross-sections remain undistorted, so that the strain at any point in the bar is equal to the radius. Therefore, at full preset, when the strain at the surface is OE, E being unity, a point at 0.6 radius would only be under 0.6 the strain and its stress would be at point I on curve OAI. The curve OAI then shows the true stresses under full preset windup, plotted against the radius as abscissa. A linear stress distribution would give stresses along line OI and excess of stresses above this line would account for the remaining torque increase.

EB indicates the nominal stress at the surface, under full windup. As the windup is released, this nominal stress becomes zero and true stress is reduced by stress EB. A trapped stress equal to EB minus EI remains at the surface. This reasoning, applied to points below the surface indicates that trapped stress at any radius are shown by differences between curve OAI and the straight line OB.

The prestressing operation also serves as an excellent inspection procedure in disclosing laps or seams in the steel, or improper heat-treatment. The latter would show up in inconsistent or inadequate spring return, following the preset load application.

The preset angle must be very closely controlled in order that consistent end indexing, which controls vehicle height, may be held within satisfactory limits. Obviously, the bars become right and left hand. Because of the hex offset, however, improper installation becomes practically impossible.

The torsion bars are protected against corrosion by three separate coats of paint. After the bars are preset at the manufacturing plant, they are washed and dipped in a phosphoric acid rinse. Bars are then dried and dipped in an epoxy black primer and baked for 15 min at 400 F. They are inverted and given a second coat of paint and again baked. At the assembly plant, a coating of rubber type paint is sprayed on, with particular attention paid to the socket ends.

The variable wheel rate feature which softens shocks in a shackled leaf spring is inherent in the torsion bar arrangement, while absent in most independent front suspensions of coil spring design (Fig. 20).

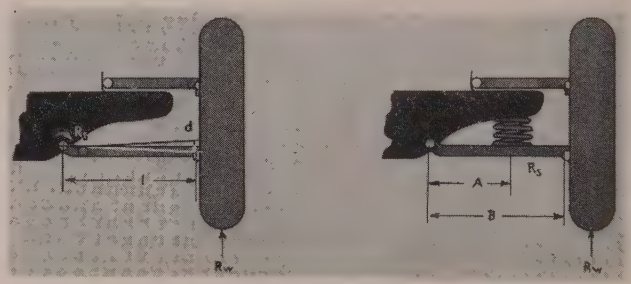
Variable rate is provided by the harmonic displacement of the control arm which changes the constant rate of the torsion bar into sinusoidal wheel rate at the ground.

Wheel Rates — The basic reason for the adoption of the independent front suspension was the opportunity to utilize softer springs with resultant lower wheel rates. This has been accomplished to a remarkable degree. More amazingly, since static deflections are neither tremendously increased, nor are the carrying capacities lessened.

Comparing the 1959 models to their 1960 counter-

Table 2 — Front Wheel Deflection Rates, lb/in.

Nominal Vehicle Rating	1959	1960
½ Ton	343	120
¾ Ton	343	120
1 Ton	463	170
1½ Ton	567	230
2 Ton	578	322
	726	310
	918	381
	619	322
2½ Ton	828	381
	1100	412
	1390	548
	1250	588
	1390	725



Torsion Bar
where: R_s = Rate, in.-lb/deg
l = Length of torque arm
d = Angle deg to give 1-in. deflection

Coil Spring
When control arms are parallel:

$$R_w = \frac{R_s}{\left(\frac{B}{A}\right)^2}$$

Fig. 21 — Wheel rate calculations

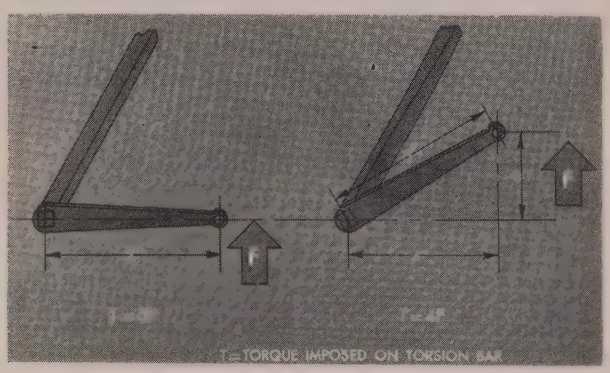


Fig. 20 — Torsion bar variable principle

Fig. 22 — Shock absorbers

parts (Table 2), the wheel rates are approximately 50% lower, contributing to outstanding improvement in vehicle ride. For example, they are reduced from 343 lb per in. to 120 on the ½- and ¾-ton models. Wheel rate decreases on comparable 2-ton models include reductions from 578 to 322 lb per in.

For reference purposes, shown in Fig. 21 is the method of calculating wheel deflection rates with both the coil and torsion bar springs.

Shock Absorbers — Another important component to be considered in an independent front suspension is shock absorbers (Fig. 22). These take on additional importance with the extremely low spring and wheel rates and reduction of friction.

The lower frequencies but higher amplitudes, with no interleaf friction damping make front shock absorbers mandatory on any and all models. The same basic shock absorber sizes available previously, provide appropriate damping and complement the lower wheel rates with revised valving.

Ideal shock absorber positioning (within economic limitations) would match its travel with that of the wheel, in a ratio of 1/1. This was impossible from the standpoint of space engineering, and a compromise had to be made. It was found that a ratio of 2/1, and revised valving, gave the necessary wheel damping characteristics as determined by the conventional methods of ride evaluation.

At present, a simple method of shock absorber mounting is used. The lower end is attached directly to the lower control arm, eliminating the need for special brackets and minimizing unsprung weight. The upper end is affixed to a bracket which is welded to the frame side rail on commercial models, and to the suspension cross-member on the heavier units.

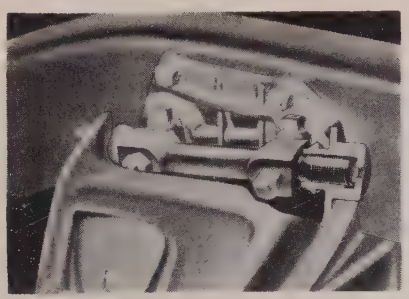


Fig. 23 — Pivot shaft and bushings

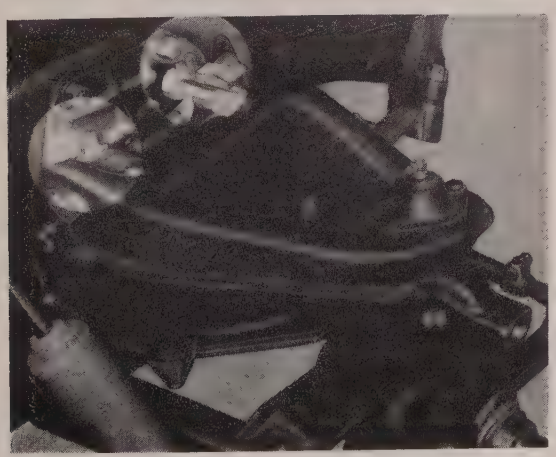


Fig. 24 — Camber and caster adjustment

Roll Rates — It was found that with an independent front suspension, vehicle roll rate is also greatly improved. A conventional I-beam suspension has the roll center located in a plane through the spring eyes. This creates a moment arm of a certain length, extending up to the vehicle center of gravity. An independent front suspension relocates the roll center to a point at approximately ground level. The moment arm length in our design is nearly the same length since the vehicle center of gravity is also somewhat lowered. The roll angle varies directly as the spring rate, which being lower, is another factor to aggravate the situation.

The single redeeming quality in this design is the effective spring center. In an I-beam suspension the spring center width is the distance between the centerlines of the springs. Independent front suspension, on the other hand, through the linkage geometry projects the effective spring center width out to the tire treads.

Mathematically, it can be determined that the roll angle varies directly with centrifugal force, and inversely, as the spring rate and as the square of the spring center distance.

During the period of development, three different commercial front suspensions were designed and labeled Mark I, Mark II, and Mark III, respectively.

Mark I was the first experimental design and incorporated rubber control arm bushings and a compression type lower ball stud. In this case, the rubber bushings and upper control arm proved lacking in anticipated durability requirements.

With these considerations, Mark II was developed, and incorporated heavier upper control arms, a tension-type lower ball stud, and larger rubber control arm bushings. The brake reaction tests performed in the laboratory, however, indicated excessive control arm movement as a result of the rubber bushing deflection. This version also had steering wheel fight and lacked component durability due to braking.

This data led to the development of the Mark III version which ultimately became the production prototype. It was found that rubber bushings, in order to provide the necessary rigidity, would have to be prohibitive in size. Threaded steel bushings, which were adopted provide all the necessary rigidity, and demonstrate a reduction in shock transfer.

Pivot Shafts (commercial) — The lower pivot shafts are bolted to the suspension crossmember (Fig. 23). The ends are threaded to serve as bearing surfaces for the control arm bushing. An interesting sidelight to the pivot shafts is their manufacturing process.

The threaded portion of the shaft is flame hardened rather than carbonitrided or case hardened, as are all other pivot shafts. A much more durable bearing surface is obtained, since flame hardening provides a surface hardness of about Rockwell 60 whereas carbonitriding draws out at Rockwell 45.

An additional factor involved is that a flame-hardened shaft will bend a few degrees farther than a carbonitrided unit, resulting in nearly 10 times greater impact resistance. Thus, a more economical and safer manufacturing process resulted in a more durable component.

Geometry Adjustment (commercial) — The attachment of the upper pivot shafts also provides a

very convenient means of caster and camber adjustment (Fig. 24). Here again, a new method is employed for which a patent application has been made. Both caster and camber adjustment is accomplished by shimming out the upper pivot shaft, an equal number of shims on each bolt for camber, a greater number of shims on one bolt for caster. An uneven distribution of shims, however, puts the pivot shaft at an angle to its attaching bracket, with resultant loads on the attaching bolts, and introduces edge contact of mating surfaces. The attachment includes a bolt which has a spherical seat, and a mating seat on the attaching bracket. A cylindrical segment spacer between the pivot shaft and bracket maintains a constant parallel relationship between the shims and the bracket, and eliminates the edge contact in clamping the shims. As a result, the bolt, instead of bending under torque, tilts in relation to the bracket, maintaining an even pressure on all shims. Since the bolt is never bent, it permits a precise degree of torque to be applied in assembly.

Antidive — The commercial front suspension design incorporates built-in correction for vehicle front end dive on braking. Primary requirements of dive control are spherical joints at the steering knuckles which allow inner pivots of the upper control arms to be tilted downward at the rear in the side view.

Basically, dive control on braking is proportional to the angle of this tilt, the front braking effort and the load transfer to the front suspension. A control arm slope great enough to provide 100% correction introduces too much caster change during suspension motion, so a compromise is necessary. The Chevrolet design incorporates a 10-deg slope which provides approximately 45% brake dive control.

Dive control, in function, is nothing more than the introduction of a rotational force from the vehicle front end dive to the spindle, in the direction opposite to that of the force imposed by braking. These forces thus tend to neutralize each other.

As shown in Fig. 25, when brakes are applied, rotational force is transmitted to the spindle support by the friction between the drums and the brake shoes and their attaching components. At the same time, there is a momentary and instantaneous load transfer to the front of the vehicle causing loading of the front end. As the front suspension is deflected downward, the slanted upper control arm displaces the upper spherical joint, and consequently the spindle support, in the direction opposite to that introduced by the initial braking force. Thus, the load transfer and the antidive are in opposition and if they were equal would neutralize each other.

Control Arms and Torsion Bar Anchors — The commercial vehicle lower arm (Fig. 26) is a 2-piece stamping, riveted to the forged spherical joint housing on one end and hinged on threaded bushings to the pivot shaft on the other. The lower ball joint

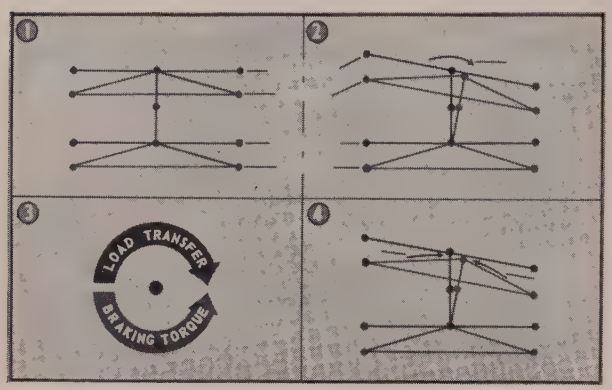


Fig. 25 — Antidive principle, light-duty models

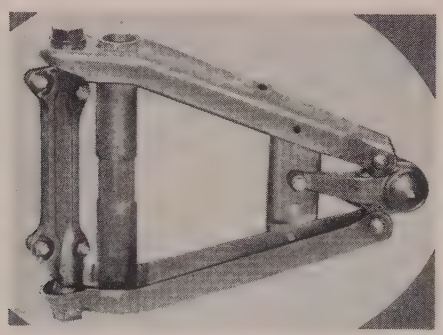


Fig. 26 — Lower control arm

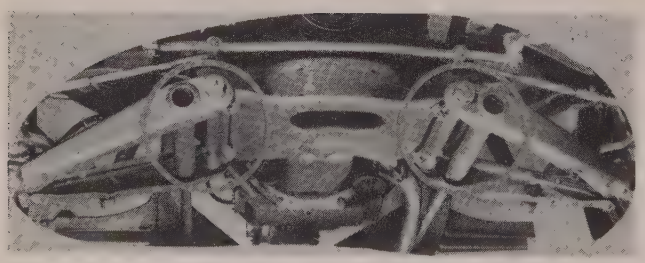


Fig. 27 — Torsion bar tube

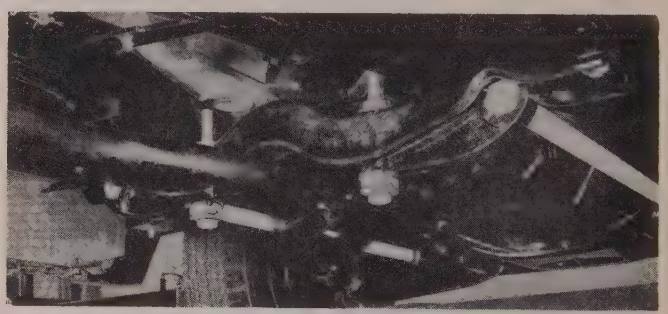


Fig. 28 — Torsion bar rear anchor

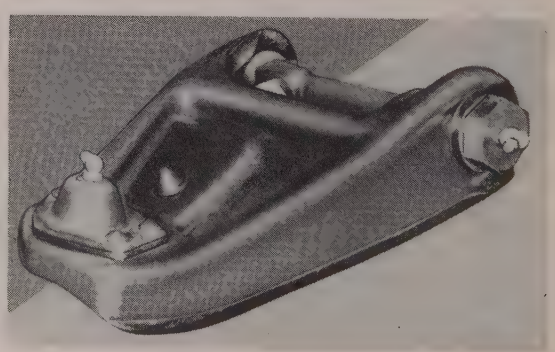


Fig. 29 — Upper control arm

supports the front sprung load in tension and also absorbs bump, side, braking, as well as shock absorber loads. A cross-brace, riveted to each section of the control arm, gives additional rigidity to the ball stud attachment and also serves as a mounting surface for the compression bumper. The channel sections of the control arms face each other. This positioning permits attachment of the ball joint assembly in a strong, simple, and low cost manner.

The torsion bar attachment consists of a tube welded to each arm of the A-frame, parallel to the pivot shaft. A hexagonal section, necessary to anchor the torsion bar end, is formed by upset forging on the 1½-ton units, and swaging on the other commercial vehicles.

Swaging of the torsion tube is a cold process, wherein the center section of the tube is simply pressed into the desired hexagonal shape.

Upset forging, on the other hand, is a hot process, capable of changing wall thickness. Beginning with a tube of a constant ¼-in. wall thickness, a hex mandrel is inserted inside the tube and a die is locked around the outside center area. The center section is then heated and the ends of the tube are compressed. This process not only forms the hex section but also increases the wall thickness in the hex area to approximately ⅜ in. With this process, the tube section can be proportional to stress.

Laboratory tests of first designs indicated that a critical tolerance existed between the torsion bar hexagonal section and the torsion bar tube. Too great a tolerance would then provide a loose fit where perhaps only two corners of the hex section would be carrying most of the load and thus reduce the life of the tube by as much as 60%. Closely held tolerances result in a tight mating fit across all six flats, with the desired component life.

The lower control arm of commercial units is of this particular design for specific reasons. Anti-dive, and the slanted upper control arms, were found to be desirable features. The low configuration of the commercial models also prescribed that the torsion bar be in the lower control arm.

An effort was made to anchor the control arm end of the torsion bar at the control arm pivot point, since this would provide only rotational movement of the torsion bar. The very large pivot bearings required to clear the torsion bar reduced precious ground clearance.

In the production design, the torsion bar is anchored to the control arm 2½ in. from the pivot shaft (Fig. 27). However, with a total spindle movement of 8¼ in. from full jounce to full rebound, the front end of the torsion bar is displaced vertically only about ½ in. in either direction.

The rear of the torsion bar is anchored through adjusting arms to a floating cross-member (Fig. 28), which, in turn, is attached to the frame through rubber biscuits and absorbs the torsional motions without transmitting them to the frame.

The floating cross-member reduced wheel fight due to frame deflection and eliminated suspension shudder at wheel hop frequencies.

Considerable development went into this torsion bar rear anchor. The first design provided for a fixed support. This, however, did not allow for the slight variations in torsion bar indexing, which together with the low rates resulted in differences in vehicle trim from side to side. In addition, fixed anchors had enough rocking motion resulting from the front vertical movement that attaching rivets were torn from the side rails. Adjustable anchors and a floating cross-member seemed to be the only answer, but not without some additional difficulty. The vertical rocking motion of the torsion bar put bending stresses on the adjusting bolt, with resulting failures. A U-bolt attachment, which would allow the anchor some angular movement, was found to be impractical since it made adjustment with the cab in place difficult. The final solution included knife edge washers at the bolt head and nut perpendicular to the torsion bar, which would allow rocking motion of the torsion bar anchor, and thus, eliminate the possibility of any bending moments being placed on the anchor bolt.

The upper control arm of the commercial models (Fig. 29), is a 1-piece stamping, with the ball joint riveted to it at its outer extremity. This upper ball joint is not loaded by the torsion bar but does support the rebound, side, and braking loads.

Control Arms and Torsion Bar Anchors (heavy-duty) — Heavy-duty truck configurations required that the torsion bar be located in the upper control arm. This presents an entirely new set of functional characteristics because the upper ball joint and upper arms become the load bearing members (Fig. 30). The heavy front end loads of these vehicles are borne by the upper ball joint in compression. The lower ball joint is not loaded by the torsion bar but does support the compression bump, side, braking, and shock absorber loads. Basically, the spindle support and spherical joint arrangement is similar to that used on the commercial models. Space, service, and economic considerations governed both the size and location of the torsion bars and, consequently, the design of the control arms. In the heavy-duty series, the bars are anchored in the upper control arm for several reasons. If mounted in the lower arm, the location would make them vulnerable to damage from obstructions in off-the-road type operations. Also, they would pass through areas generally reserved for transmission power take-offs. Furthermore, if they were

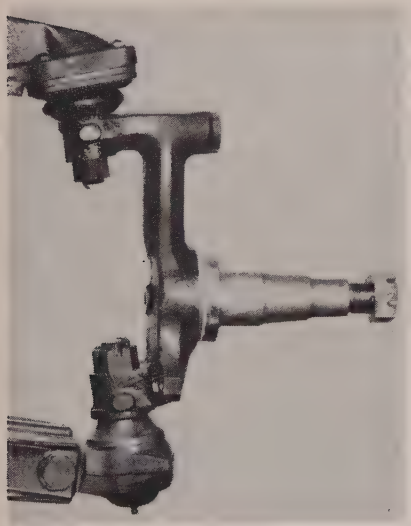


Fig. 30 — Spindle support attachment, heavy-duty

mounted on the lower arm and extended horizontally rearward, they would be far enough below the frame side rail to necessitate large brackets for the rear anchors.

The same reasons that make antidive possible and practical on commercial models, make this feature impractical on the heavier, longer wheel-base units.

In an area where wheelbases of 200 in. are not uncommon, and the larger percentage of braking effort is absorbed by the rear wheels, neither the load transfer nor the braking effort is sufficient to allow any appreciable degree of dive control with a practical upper control arm slant.

These factors, coupled with the higher ratio of spring rate to front end load on the heavier trucks, reduce the amount of dive for a given rate of deceleration. Actual proving ground tests show that with a maximum practical control arm slant of 10 deg, dive was reduced only 2.5% in 5 fps^2 decelerations and only 6.5 in 10 fps^2 decelerations, as compared to a vehicle with no antidive design.

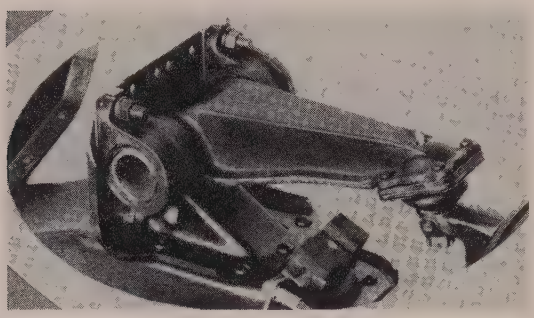


Fig. 31 — Upper control arm, heavy-duty

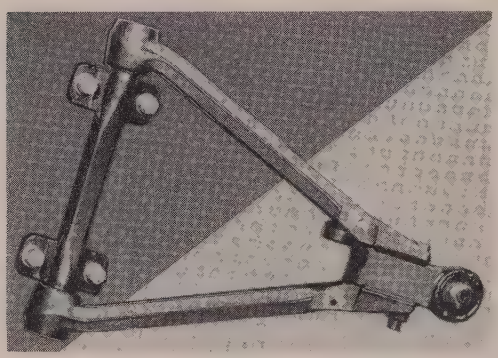


Fig. 32 — Lower control arm, heavy-duty

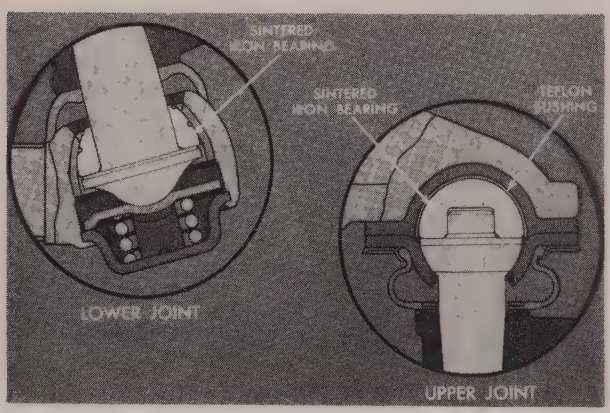


Fig. 33 — Spherical joints, heavy-duty

The upper control arm is a sturdy Armasteel casting, incorporating integral pivots (Fig. 31). To ensure equal load on each pivot bushing, the torsion bar seat hexagon is located midway between the bushing centers. Since the torsion bar must pass through the rear pivot shaft to its centered seat, the inside diameter of the pivot shafts has to be quite large. A forging would require extensive machining operations to clean out this bore and form the correct torsion bar hexagonal seat. For this reason an Armasteel casting, with a minimum yield strength of 60,000 psi was chosen as the material for the upper arm. Having a hardness range of 197–241 Brinell, it provides the required durability for the torsion bar hex seats. In addition, this material is flame hardened to the required degree on the pivots. The spherical joint is then riveted to the outer end of each control arm.

The lower arms, on the other hand, while not being the load bearing members, must absorb tremendous shock loads as the vehicle travels over road irregularities, and react to brake and side loads. The control arm unit consists of two identical forgings (Fig. 32), with the lower ball stud assembly dovetailed between them and secured by two "through" bolts, the outer of which incorporates a shear washer to prevent any slipping during abusive bumper or braking loads.

Load distribution to the upper and lower pivot points (Fig. 33) is entirely different from commercial models, since the torsion bar acts on the upper rather than the lower control arm. The upper joint thus, becomes the load carrying member and is in compression.

Spherical Joints (heavy-duty) — Typical of the problems encountered is the following example. It was discovered, that in the early experimental models of the 7000-lb suspension, during full-stroke laboratory and Belgian Block tests the loose sintered iron bearing of the upper stud developed a discolored and spalled surface.

An exhaustive series of tests involving varying ball sizes and the choice of other bearing materials led to the final selection of an assembly in which both the rocking and turning motions would be taken by a sintered iron bearing in a teflon-lined socket. This design not only proved superior to the metal-on-metal type since the heat buildup was reduced, but would also maintain constant friction characteristics throughout its prolonged service life.

The lower spherical joint is a tension type, more conventionally designed with a sintered iron bearing in a hardened socket and a cold upset stud. Contact areas are induction hardened. A strong spring maintains the lower ball on its seat during severe shock absorber rebound loads.

The control arm pivots (Fig. 34) were the result of extensive development and testing. As an example, both the upper and lower control arm pivots of the heavy-duty vehicles were originally designed to have the "interrupted" or "breach block" type of bearing surface. This consists of threads of zero pitch profile in two 90-deg segments, each separated by another 90-deg segment in which the threads have been machined off. This principle works similarly to the more common threaded bushing but requires no spacer bushing, as the threads are machined directly on the pivot shaft and on the

control arm bosses. This design, of course, permits easier installation of the control arms and an overall size reduction, since space is not needed for a separate bushing.

In the process of development it was found that this pivot design could not tolerate the slightest misalignment, and did not provide an adequate bearing surface. Also, the friction was great enough to have an adverse effect on the ride.

Pivot Shafts (heavy-duty) — The final upper pivot design provides phenolic bushings, reinforced with cotton fibers which contain a small amount of graphite to reduce breakaway friction. These have proved to be extremely durable, and in addition, if slight misalignments occur in assembly these do not cause binding of the bushing as might be encountered in a metal-to-metal surface.

The most satisfactory bushing design for the lower control arm pivot proved to be the full-threaded type. The bearing surface is provided by a threaded control arm end, which is turned directly onto the pivot shaft. With all requirements, such as minimum shaft sizes, loads imposed, angular deflection, and durability taken into account, rubber bushings became impractical.

Since the torsion bar load is not on this arm, these threads have almost no friction. The threaded pivots on all other suspensions presented no friction problem because of the smaller thread diameter involved.

Geometry adjustments on heavy-duty models are accomplished differently than on commercial units. Camber is adjusted by the insertion or deletion of shims between the upper control arm bracket and the mounting surface on the frame bracket.

The lower control arm provides the caster adjustment. Elongated holes in the pivot shaft ears permit a fore and aft movement of the entire control arm assembly. Lateral serrations in the attaching surfaces of the pivot shaft and its mounting surface permit predictable accuracy. Each serration changes caster angle by ½ deg.

Steering Systems, General — Since the independent suspension, by its design, creates the totally unrestricted movement of each wheel about its pivot centers, it requires a tailored steering linkage which must offer little or no restriction to this wheel movement, yet assure directional control at all times.

The original design criteria for the steering system included: low steering effort, understeer for proper handling, reduced turning radii, reduction in fight, road shock or feel at steering wheel, relocation of the steering wheel to the most comfortable position, elimination of lash in steering gear at straight ahead position, improved steering wheel return, consideration of Ackermann, elimination of adjustable wheel stops and greased-for-life bushings wherever possible. Here, a 2–1 safety factor to material yield strength was used.

Four general types of steering mechanisms can be utilized on independent front suspensions (Fig. 35). These are: center arm, Haltenberger, long and short tie rod, and parallelogram. In all cases, except parallelogram, the tie rods are of unequal or incorrect length and cannot be made to correct for suspension characteristics.

The parallelogram steering linkage was adopted for both the commercial and heavy-duty models, and provides a completely balanced system where, under identical conditions, the steering action on one wheel is the same as on the other.

Many studies were made on optimum steering linkage location. Due to space requirements, forward steer had to be adopted.

Development Systems — Initially, commercial truck steering linkage provided for the straightest components possible. The relay rod was located in front of the steering gear, so that the pitman and idler arm ends transcribed the same arcs as the steering arms. This arrangement provided the greatest ratio buildup on turns, but had to be abandoned because it forced the relay rod so far forward of the outer pivots that tie rod angularity to the centerline of the lower control arm was 20 deg. This angularity made the tie rod fight the deflection of the rubber-bushed original suspension and was a factor in its being discarded. To reduce the tie rod angularity, the relay rod had to be offset 4½ in. This offset and the fact that the steering still needed improvement resulted in another complete redesign.

The second design (Fig. 36) was the bell crank system which incorporated double-ended pitman

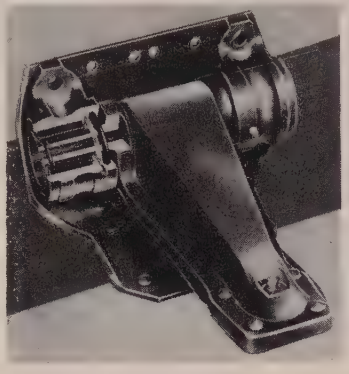


Fig. 34 — Upper pivot shaft and bushing, heavy-duty

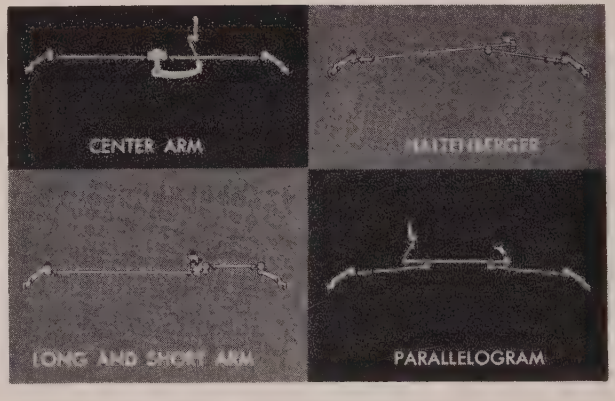


Fig. 35 — Basic types of steering systems

Fig. 36 — Bell-crank steering system

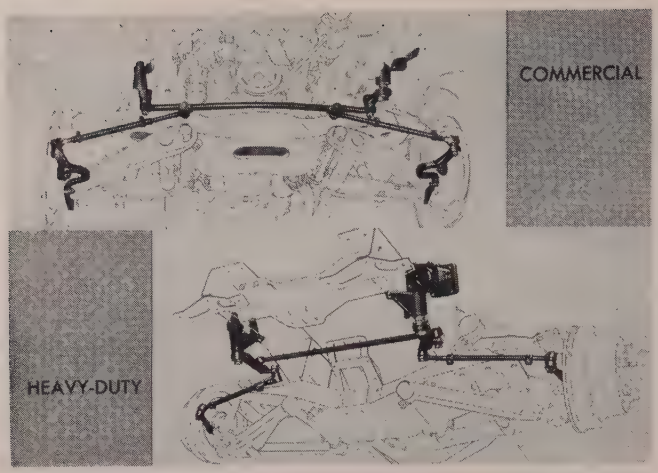


Fig. 37 — Chevrolet steering systems

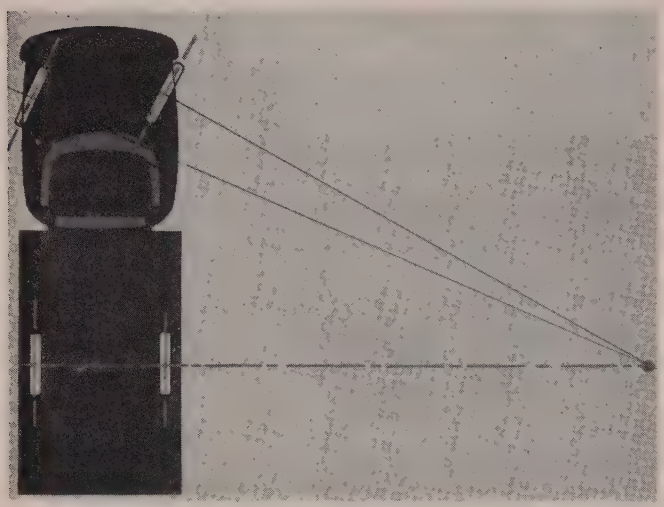


Fig. 38 — Ackermann steering layout (no tire slip angles)

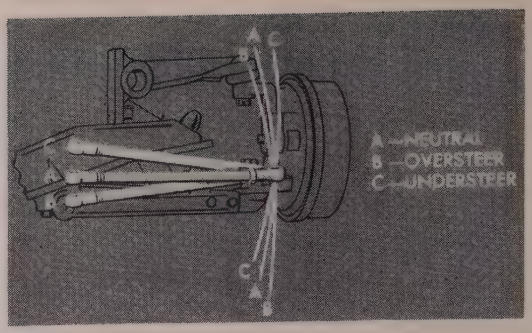


Fig. 39 — Front suspension roll steer

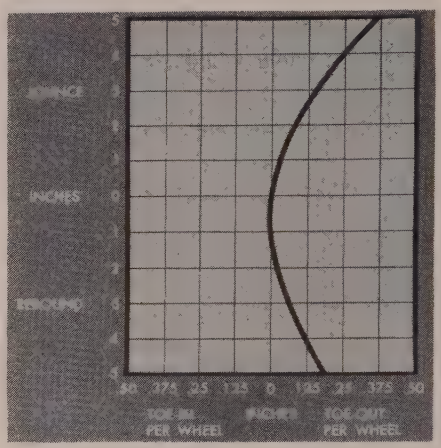


Fig. 40 — Toe change (average, all models)

and idler arms and parallel tie rods. This arrangement provided the theoretical Ackermann by locating the tie rod inner pivots inside the centerlines of the idler and pitman arm. Several vehicles were built incorporating this system. Testing, however, revealed several deficiencies. There was a marked lack of rigidity between the wheels, with some large stresses imposed on the idler bushings. Another peculiarity of the linkage system resulted in right-hand wheel wobble in a left turn, after the tires became corner worn. Further development, no doubt, would have eliminated some of these difficulties. However, since this system involved a cost penalty to begin with, it was decided to revert to the conventional parallelogram design.

At this stage the heavy-duty truck had one major steering linkage design overhaul. This was the result of space requirement changes involving steering gear positioning and did not alter its basic design appreciably.

The production commercial model steering system (Fig. 37) has the relay rod pivoted at the pitman and idler arms by nylon and rubber, greased for life bushings. Tie rods are attached to the relay rod of these bushings through conventional metal inner pivots. The tie rods extend forward at about a 4-deg angle and attach to the steering arm bosses. While tie rods, which have angular relationship to the centerline of the control arm, are undesirable, small angles are tolerable, and in this case give about 1 deg of additional toe-out in full turn.

Heavy-duty models have a similar linkage, except for the necessary differences required for strength and Ackermann. Here the pitman and idler arm have double bosses, the outer is located on the centerline of the pitman shaft and the inner is offset. The relay rod, which has metal socket assemblies, is attached to the outer bosses so that parallelogram articulation occurs. The adjustable tie rods, also made with metal socket assemblies, are located by the inner bosses.

Because of variations between the commercial and heavy-duty models, two different systems are employed (Fig. 38). With a front-mounted steering system there is some difficulty in designing for the theoretical Ackermann, unless a prohibitive wheel offset is utilized, since the angle of the steering arm is limited by the wheel.

It was found that 50–60% less than the correct condition was entirely satisfactory for the commercial trucks because of the relatively light front end loads. This decision was further supported by experience with the Chevrolet passenger car which has operated for some years with less than perfect Ackermann, yet with no complaints about tire wear, loss of handling ease, or even tire squeal. In the case of the heavy-duty trucks, it was felt that since these operate with heavier front end loads, lower speeds, larger tires, and therefore, smaller slip angles, any great departure from the theoretical Ackermann would result in tire scrub and even dam-

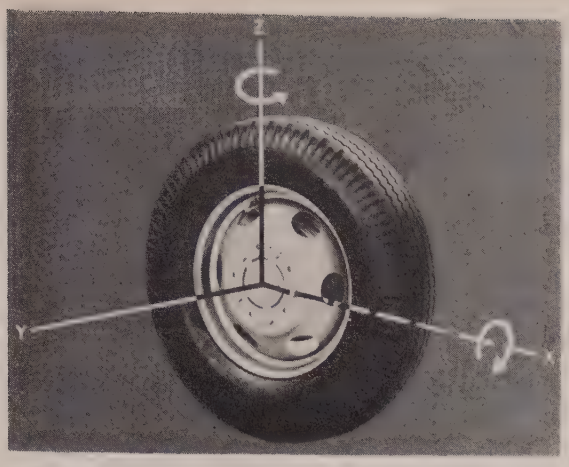


Fig. 41 — Gyroscopic effect

age to the road surface under certain conditions.

Therefore, a system had to be designed which would give the required wheel relationship in a turn. The overlapping position of the tie rods in relation to the relay rod, at the pitman arm and idler arm is such that movement of the pitman and idler arms, in turning the wheels, causes a lengthening of the inside tie rod and a foreshortening of the outside tie rod. This, of course, provides a greater wheel angle and the theoretical Ackermann can be designed from considerations of the degree of tie rod and relay link overlap.

Roll Steering — A good handling truck demands front suspension understeer, which is toe-out in jounce and toe-in in rebound. In a conventional I-beam design attempts have been made to provide front suspension understeer in roll on some poorly handling models. This was accomplished by positioning the steering gear so that the drag link sloped downward and the arc described by the end of the drag link was mismatched with the arc of the third arm ball stud. In order for it to be effective, this configuration required that the spring be shackled at the end away from the steering gear. Thus, a roll to the right caused the wheels to turn slightly to the right and provide understeer, and vice versa. However, the amount of steering effect was limited by the mismatch that the two arcs would tolerate, without objectionable wheel fight on rough roads.

Roll steer on an independent front suspension, with a parallelogram forward steering system, is obtained by locating the inner and outer pivots so that these cause the tie rod to transcribe a path which nearly overlaps the forced arc of the supension at the arm attaching boss (Fig. 39). Any mismatch will result in a roll steer characteristic and can be controlled. Another factor in controlling the amount of steer in jounce and rebound is the length of the tie rod. On these vehicles, where the tie rod is approximately 2 in. longer than theoretical, the system was designed for understeer in jounce and nearly neutral steer in rebound.

$$\text{Per cent steer} = \frac{\text{Toe change per wheel}}{\text{Vertical wheel movement}} \times 100$$

The Chevrolet commercial configuration, (Fig. 40) with a maximum jounce of 4½ in., moves each wheel almost a half inch to the toe-out position, whereas a maximum rebound of 3.75 in. results in little or no toe change. Thus, with the vehicle in a roll, approximately 5–6% understeer is provided.

The amount of understeer was varied in the development program from 2 to 17%. No serious variation in handling could be determined, but extreme toe changes at higher percentages made adherence to the safer, low limit of 5 and 6% the obvious choice.

The heavy-duty suspension is designed for a maximum jounce and rebound of 5 in. each. Here the steering system is designed for approximately a 3% understeer. Maximum jounce of 5 in. results in 0.32 in. of toe-out, while a rebound of 5 in. provides 0.18 in. of toe-out.

Accurate initial toe-in settings are required to allow the vehicle to properly respond to the designed geometry.

A most important consideration in steering systems is that they have adequate strength. Strain gage readings were obtained on each part of the steering system of all models. From this data adequate sizes and materials were selected.

As has been previously stated, the good handling truck demands understeer. It is indeed fortunate that the linkage needed for proper handling requires the same geometry as for no wheel fight. One of the greatest contributors to wheel fight is the precessional force resulting from the angular spindle displacement of the front wheel which is, in essence, a gyroscope (Fig. 41). The decision to adopt an independent front suspension reduces this problem.

Gyroscopic Effect — A front wheel, after hitting a chuck hole, is toed in by the impact, which causes a vertical precessional force which lifts the wheel. When an I-beam axle is used, the wheels are directly coupled so that the opposite wheel is forced into the ground and toed-in at the same time. The normal aligning forces act on the opposite wheel, causing rotation to toe-out. Processional forces lift this wheel. The interaction between wheels becomes self-sustaining and the shimmy cycle is established. This shimmy causes the most severe wheel fight known to the automotive engineer. Unfortunately, steering linkage adjustments could not be made through a single drag link on a solid axle, to correct for this complex interwheeel action. Some steering linkages undoubtedly aggravate this action.

An independent suspension reduces the shimmy problem by uncoupling the front wheels, or removing the rigid, low mass I-beam. The vertical precessional forces now act against the mass of the vehicle and the torsionally stiff front end of the frame. Shimmy is never completely eliminated, however, and if not compensated for by proper linkage design, can result in wheel fight of disagreeable proportions. The understeering linkage, since it toes out in jounce, will complement the individual wheel motion caused by gyroscopic forces and not produce wheel fight.

Wheel shimmy is not new. Maurice Olley has shown the results of sustained shimmy as it existed on the I-beam axle passenger cars, as far back as 1927.

An unusual situation develops in the gyroscopic action which occurs during parallel hop and tramp of an independent front suspension, as compared to an I-beam unit. Gyroscopic action does not

occur on an I-beam front suspension in parallel hop since there is no camber, or significant toe change, when both wheels simultaneously move up or down an equal distance. An independent suspension involves gyroscopic precession because both jounce and rebound are accompanied by camber and toe changes. The forces however, are balanced and, therefore, absorbed in the linkage.

With an I-beam suspension in tramp, the worst possible gyroscopic action occurs since the axes of both wheels are displaced in the same direction to compound the effect. An independent suspension in tramp dampens this precessional force as both jounce and rebound produce negative camber, to provide partially balanced forces. In addition, the linkage in jounce produces toe-out, which is in the opposite direction from the precessional force resulting from camber change.

Steering Gear Positioning — In addition to providing correct toe-in adjustment, an inherent characteristic of a steering system with two tie rods is its ability to allow for steering gear on-center positioning, in the straight ahead direction.

Elimination of lash for all positions of the steering gear would be highly uneconomical. Steering gears are, therefore, designed to have zero lash on the high point, which normally should be the straight ahead position. However, with a single adjustable tie rod, correct toe-in may take the steering gear off this high point when the vehicle is going straight ahead. Two tie rods permit this no-lash adjustment, eliminating the wandering feel in the straight ahead direction.

Still another consideration involved reductions in turning radii, through optimum placement of components. As an example of the accomplishment of this objective the final version of the ½ and 2 ton trucks have turn radii 1.6 feet smaller than their 1959 counterparts, without adjustable wheel stops.

An attempt was also made to utilize all of the steering gear travel available in its design, in order that the highest overall ratio could be obtained which would result in the minimum possible turning efforts. Consequently, the overall ratios were increased over 1959 by approximately 10 percent, to 28.7 to 1 and 32.6 to 1 for commercial and heavy-duty trucks, respectively.

Finally, as a general steering system design criterion, all linkage members were to be made as straight as possible to minimize stress and to maintain maximum linkage rigidity between front wheels.

Steering gear assemblies are of the conventional design incorporating the recirculating ball-nut. Only two ratios, 24 to 1 and 28.14 to 1, are used for commercial and heavy-duty trucks respectively. Pitman shaft diameters, one measure of gear size, are 1.12 and 1.38 inches. These diameters correspond, almost exactly, to previous design. Steering wheel diameters have been reduced one inch so that they are 17 inches for commercial models and 19 inches for heavy-duty trucks. These new wheel diameters were in accord with general findings relative to driver comfort and do not impair steering effort because of benefits derived from ball type steering knuckle pivots.

The driver comfort program was conducted in 1957, using a vehicle with an adjustable steering wheel. A number of experts, ranging from the chief truck engineer to independent truck drivers, expressed their opinion on optimum location after hours of driving time for each. These tests established the 1960 wheel position, and it was decided to adhere to these without regard for steering gear location. This becomes especially difficult where the same steering gear and linkage is used for conventional and low cab forward (LCF) models.

To overcome this difficulty two universal joints are used. These allow the optimum upper steering shaft positioning and the use of a short mast jacket, completely clamped in the cab. Further advantages of two joints include the isolation from the road shock and wheel fight occasioned by relative movements between the cab and a frame-mounted gear. The commercial trucks utilize a pot-type universal joint at the upper end for angularity and a flanged coupling at the lower where there is no shaft misalignment. Practical experience indicates that there are benefits derived from the use of two joints, even with straight shafting.

The heavy-duty trucks have three separate shaft arrangements, for conventional cab, school bus and LCF models. All of these use pot-type joints at either end because of the greater angularity encountered.

Both the light- and heavy-duty steering systems are designed to be completely compatible with, and multiply the benefits of, the independent front as well as the rear suspensions.

Rear Suspension, general — A drastic redesign of the front suspension necessitates a complete reevaluation of the rear springing systems as well. Front improvements alone would not eliminate the obstacles standing in the way of a better ride. It is impossible to change a front suspension without involving the rear, or the entire vehicle for that matter. The ultimate result is not specifically due to one or the other but is the sum total of the vehicle as a whole.

The ride frequency of the rear suspension must be compatible with that of the front. Steer characteristics must also complement each other. It is impossible, for example, to correct rear suspension oversteer characteristics by an excessive amount of understeer at the front, and vice versa. Still other factors to be considered are roll rates. The excessive differences in rates, especially if the suspensions are separated by a torsionally stiff frame, can nullify any handling improvements that could be attributed to the suspensions alone.

The most popular rear suspension spring in truck use today is the leaf spring. Benefits derived from its characteristics are well-known (Fig. 42). In addition to their function as elastic members, leaf springs do a fair job of maintaining axle alignment as well as perform, to a degree, as shock absorbers by way of interleaf friction.

On the opposite side of the ledger, however, is the variation in interleaf friction and inefficient use of material, resulting in excessive weight. As the designer makes the leaf spring softer, the capacity of the spring to serve as a locating member is impaired. This is not to say that disadvantages of rear leaf springs outweigh the benefits to a point where their usage would be incompatible with independent front suspension. A great deal can be done and has been done to minimize the effects of the unfavorable qualities of leaf springs, to make

them compatible with independent front suspensions.

The area of rear suspension design is one of extreme compromise. Whereas, for an entire line of vehicles, front end loading, in going from curb weight to maximum gvw, may only vary from 2000 lb on the smallest to about 6000 lb on the largest, the rear suspension loading on a single axle may range from 1400 to 22,000 lb.

Because of this variation, it is necessary to divide a complete line of vehicles into several rear suspension design categories to obtain the proper compromise of ride and carrying capacity.

Rear suspensions, in the planning stage, also involved many different designs. These included 4-link, 2-link, trailing arm, parallelogram linkage, rigid axle, swing axle, and others.

Here it was decided to adopt three different suspension designs, each tailored to the vehicle type. Rigid axles are retained throughout, and a 2-link system is used on the lighter units, a conventional leaf spring on the intermediate, and a variable rate suspension on the heavy-duty.

Two-Link Rear Suspension — On commercial vehicles it is evident that rates can be low, approaching those of passenger cars, and still maintain their load-carrying capacity. Because of the low rates, a frictionless suspension system, with the spring used as an elastic member only, has proved most satisfactory. The following reasons governed this selection:

1. The rates are sufficiently low so that interleaf friction of a leaf-type spring suspension would be extremely apparent in ride.
2. A leaf spring of correspondingly low rate would not be adequate as an axle locating device.
3. The linkage permits the outward movement of the effective spring centers so that the lower rate does not induce excessive roll.
4. The load range involved does not lead to impractical spring sizes from the standpoint of manufacturing.
5. Springs can be used which will store approximately four times as much energy per pound of steel as leaf springs.
6. As with the independent front suspension, the rear linkage system permits a controlled suspension geometry. Thus, a precise degree of understeer can be designed into the system to offset any tendency for vehicle rear oversteer due to larger slip angles of loaded tires.
7. A more nearly constant relationship can be maintained between the rear axle pinion and the propeller shaft, minimizing the unfavorable angles through which the universal joint has to transmit driving forces under axle windup.

The linkage type rear suspension, provided for the ½- and ¾-ton models (Fig. 43), has maximum capacities of 3500 and 5200 lb, respectively. Both are identical from the standpoint of design, with size of components proportional to capacity. While a rigid axle is retained, a link system incorporating coil springs, provides an improvement in ride and handling, corresponding to that of the independent front suspension.

Control Arms — Two trailing arms, to which the rear axle is rigidly attached, control all driving and braking torque. Low rate coil springs are cradled between the arms and the frame sidemembers, while a rear tie rod controls tranverse or lateral forces.

The control arms are constructed of two identical channel sections, attached back to back, which are connected to the frame through rubber bushings to provide maintenance-free operation. The point of attachment consists of separate cross-braces extending from the frame cross-member to the side rails. The pivot bracket is attached to the cross-brace, as near the centerline of the vehicles as the cross-member will permit. The control arms then extend diagonally rearward and are bolted beneath the axle. Shock absorber brackets are attached at the control arm rearmost extremity.

The trailing arms alone control the driving and braking torques transmitted by the rear axle. Consequently, the length of the arms becomes important. Short control arms not only magnify axle reactions, resulting in excessive rear end squat and lift, but also have an adverse effect on bushing life. In addition, tests indicate that longer arms elimi-

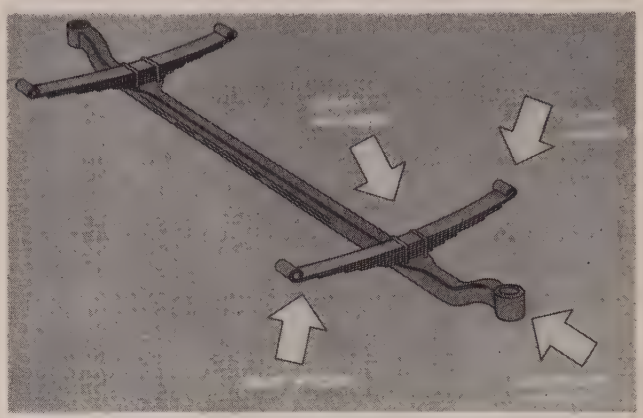


Fig. 42 — Leaf spring functions

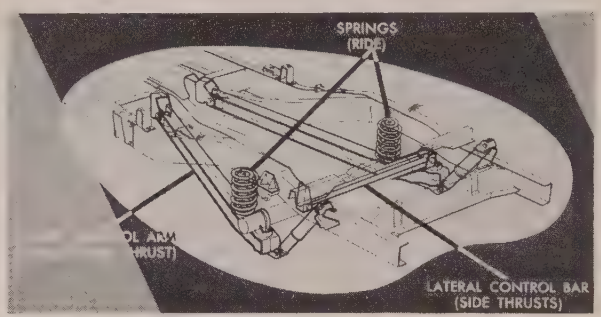


Fig. 43 — Rear suspension (½- and ¾-ton models)

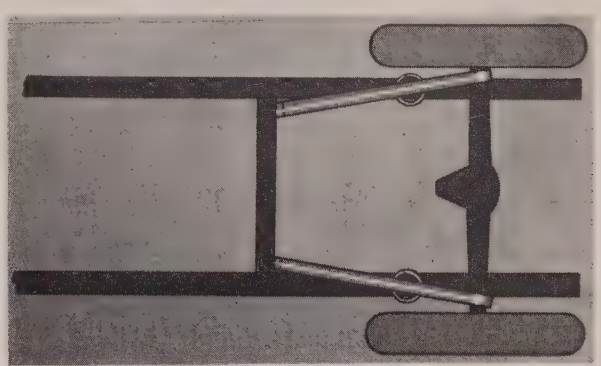


Fig. 44 — Effective spring centers

nate brake and forward power hop and minimize reverse power hop.

The angle of the control arm in relation to the axle is important from another aspect (Fig. 44). Effective spring centers in this design are determined by the intersection of a line, which extends from the control arm pivot point, through the coil spring centers, and the axle centerline. Roll stability is thus enhanced since the spring centers are wider than those obtainable with a conventional leaf spring suspension.

Another factor to be considered in control arm positioning (Fig. 45) is the amount of twist transmitted to the arm and bushing by differential wheel movement. Since the pin is positioned normal to the control arm, the inboard positioning of the pivot point imposes less angular twist on the control arm than if the control arms were parallel to each other from their points of attachment to the axle.

The twist transmitted to the arm and bushing by differential wheel movement led to a development program as to the type of control arm construction which should be used in production. Considered were hat sections with a bottom plate, tubular sections upset to a rectangular form under the axle and spring center, and fabricated I-beam weldment sections. From the standpoint of both economics and durability it was determined, through extensive laboratory and vehicular testing, that the channel back to back was the most appropriate design. The prime considerations, in addition to beam strength, included the torsional rigidity of the arm, its effect on the vehicle roll, as well as the durability of the axle and bushing.

Rear Roll Steer — Rear suspension roll steer of linkage systems can be predicted graphically by determining the roll axis of the rear axle (Fig. 46). The roll axis is a hypothetical line joining the forward and rear centers of roll of the rear suspension. The forward center is that point at which the control arms intersect if projected forward. This point would lie in the median plane. The rear center is the point at which the lateral control arm crosses the median plane of the vehicle. The line joining these centers is the roll axis of the axle and its angle from the horizontal is the steer angle of the axle at this particular standing height.

If the rear center is above the forward center, understeer exists. Oversteer exists if the rear is below the forward center. The tangent of this angle, times 100, is often used as an expression of percentage of steer.

As the standing height changes, the relative height of the points change. If the vehicle is loaded, it may be found that the forward point goes down more than the rearward point. Thus, the slope of the roll axis changes and the axle has less oversteer, or more understeer at higher loads.

In practice, roll steering of a link rear suspension is accomplished by the lengthening and shortening of the effective lever arm about which the wheel tends to pivot in jounce or rebound (Fig. 47). With a vehicle roll to the right, the outside or right rear wheel will go into jounce, while the left goes into rebound. Assuming the roll axis of the axle slopes upward to the rear, a roll to the right decreases the effective length of the right-hand control arm. The opposite effect is true of the left-hand side where the effective control arm length is increased. Thus, the axle is rotated to the left with relation to the median plane of the vehicle, providing understeer.

Oversteer, of course, requires a roll axis of the axle sloping downward to the rear, where the opposite effective control arm lengthening and shortening will take place.

The degree of rear suspension roll steer required is entirely dependent on the overall vehicle characteristics, since it involves many variables. Included are the height of the center of gravity and tire loading, both of which vary with the load, tire characteristics, location of the c.g., the neutral steer point, and the vehicle roll center. For this reason, suspension steering is not an analytical, but rather a dynamic development, requiring many miles of test driving under varying conditions before the optimum handling characteristics are achieved. In the case of the Chevrolet design, 3% understeer, loaded, proved ideal.

Lateral Control Bar — A 2-link rear suspension system also requires a tie rod (Fig. 48) to maintain lateral relationship between the axle and the frame.

Table 3 — Rear Wheel Rates, lb/in.

Vehicle	1959	1960	Vehicle	1959	1960
½ Ton	200/270	199	1½ Ton	746/934	1375
	284	261		934/1394	1375/3408
¾ Ton	290/436	299	2 Ton	1442/3475	540/2380
	501	425		1713/4222	750/2500
1 Ton	496	365/495	2½ Ton	1552/3462	750/2500
	550/750	497/1250		1880/4100	900/3100
				1960/4330	975/3750

Fig. 45 — Control arm pivot

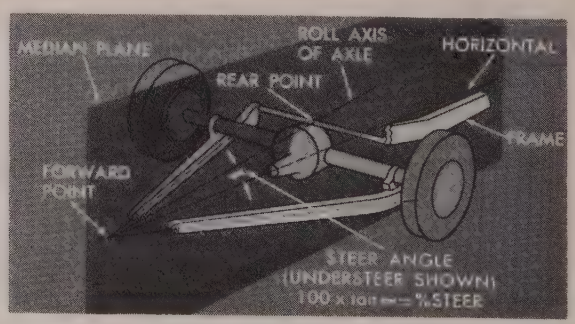


Fig. 46 — Roll steering

The solid, 1-piece steel bar is attached through rubber bushings to brackets welded on the left to the sidemember, and on the right to the axle housing. Also, as stated previously, since the lateral control bar determines the location of the roll center of the rear suspension, it is important in establishing the steer characteristics.

Considerable development led to the choice of the 1-piece rear axle tie rod. This rod consists of a solid steel bar, with eyes resistance welded at each end. In the process of development many types of construction were considered, including tubular rods with welded on eyes, 1-piece forgings, forged eyes welded to a solid rod, and forged eyes welded to a tube. Also, a variety of welding techniques was studied for each application.

Final specifications for quality control of these bars require that the ½-ton rod be capable of withstanding 10,000 lb in tension or compression and that the ¾-ton tie rod be capable of withstanding 15,000 lb. Each axle tie rod produced is proof tested to 6000 tensile load to assure that no inadequately welded pieces are incorporated in vehicles. It is interesting to note that calculated stresses on this component did not even approach the actual strain gage measurements.

Spring positioning ahead of the axle places it at a slight mechanical disadvantage (Fig. 49). Thus, capacities of the spring at the pad are meaningless and reference from this point on is to wheel rates.

Coil Springs — A small measure of how well the job of providing softer springs was accomplished is gained from a wheel rate comparison between 1959 and 1960 models (Table 3). The wheel rate for a 1959 leaf spring ½-ton rear suspension was 200 lb/in. for the first stage and 270 lb/in. for the second stage. In comparison, the 1960 rate is a constant 199 lb/in. Optional springs are reduced from 284 to 261 lb/in.

Three-quarter-ton models provide similar reductions, from 290 lb/in. for the first stage and 436 lb/in. for the second stage, to a constant 299 lb/in. Optional springs are reduced from 501 to 425 lb/in.

The important fact is that these are pure rates not involving friction. The elimination of friction with the link rear suspension greatly improves the ride, with no major decrease in rates, yet with increased gvw capacity.

The springs used are of high alloy steel wire 0.65 in. in diameter on base ½-ton models, with an inside coil diameter of 4.63 in., approximately 7.61 total and 6.11 active coils. The ¾-ton base vehicles utilize larger springs of 0.785 in. wire, an inside coil diameter of 5.20 in., approximately 7.58 total and 6.03 active coils.

Utilization of pigtailed ends on the coil springs (Fig. 50) eliminates any necessity for machining the ends of the coil flat to conform with a mounting surface, or the necessity for holding end relationships to fit into special mounting pads. This design is also more economical in steel usage since the amount of inactive wire is less. Another tremendous advantage is that all of the inactive wire will fold internally and reduce the solid spring height.

Rear Suspension, 1 and 1½ ton — The only workable steel suspensions, which adjust rate to maintain constant effective deflection, accomplish this

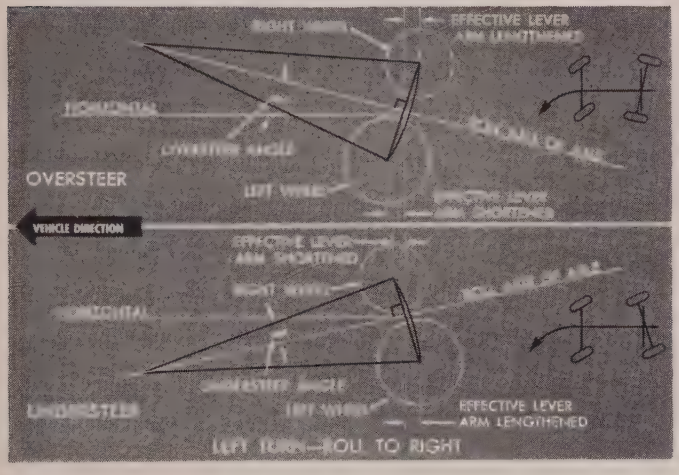


Fig. 47 — Roll steer

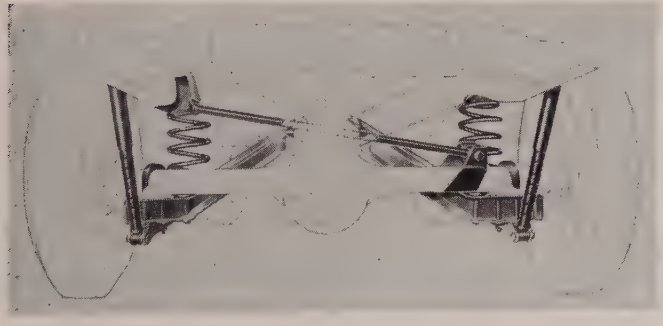


Fig. 48 — Lateral control bar

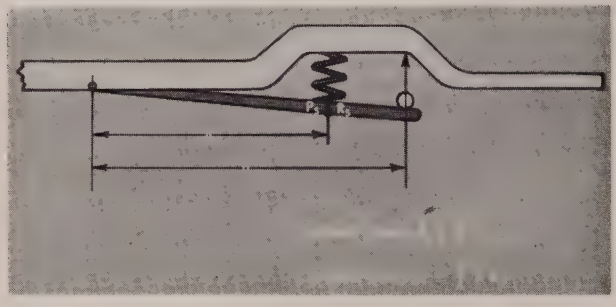


Fig. 49 — Rear suspension rate

Fig. 50 — Rear suspension coil spring

by using progressively less spring material as the load increases. The springs of the 2- and 2½-ton models were considered for the 1- and 1½-ton models but because the payloads were not large enough to utilize the full benefits, these were not adopted. Therefore, it became evident that a middle-of-the-road design with 2½-in. leaf springs and conventional hangers and shackles would be most appropriate at this time (Fig. 51). The entire design, however, was overhauled with notable improvements to hangers, shackles, and leaf stresses, making frequencies and rates compatible with the front suspension.

The use of auxiliary springs, was found unnecessary to achieve the rate variation for loads in this range, this being much more satisfactorily accomplished with the use of 2-stage springs.

The greater camber of the upper several leaves provides highly acceptable frequencies and amplitudes of the rear end ride unloaded. Spring stiffness is also sufficiently high to provide adequate axle alignment and, of course, eliminates auxiliary spring hangers and their unpleasant slapping, as well as excessive and inefficient spring steel usage.

Springs for both the 1- and 1½-ton uints are 52 × 2½ in., with eight leaf standard and 10 leaf optional on the 1-ton, and 10 leaf standard and 12 leaf optional for the 1½-ton vehicles. Springs are attached to the axles through U-bolts and durable Armasteel anchor plates. The spring is pinned at the front to a conventional malleable iron hanger which is different from 1959 models only to conform to the new frame configuration.

The rear hanger incorporates a new tension-type shackle which is the key to optimum rear suspension steer characteristics. Oversteer was always a problem in this category, especially where comparatively heavy loads are carried on single tires.

Assuming similar positioning of the rear hangers, the choice of a tension- rather than a compression-type shackle is a classical but effective method of providing rear suspension understeer.

Roll Steering — Understeer may be added to the rear suspension (Fig. 52) by positioning the spring so that the fore and aft centerline through the spring eyes is sloped down toward the front of the vehicle. This geometry produces the same changes in the effective lever arms about the axle, as do the control arms on a 2-link suspension. In a turn to the left, the vehicle rolls to the right and slope of the radius of action on the right-hand spring is increased, while that of the left-hand spring is decreased. This causes the axle to steer to the left, opposite to the direction of roll and the vehicle tends to steer out of the turn, or understeers.

Changes in rear wheel rates are quite spectacular in the 1- and 1½-ton series (Table 3). As an example, the 1½-ton models had a rate of 1375, as compared to the 1960 2-stage rate of 764 and 934 lb/in. Compatibility with the rates of the front suspensions of these vehicles is maintained to provide the optimum in ride and handling characteristics.

Rear Suspension, 2- and 2½-ton — Meanwhile, on the heavy-duty vehicles, it was evident that some method of increasing the rate variation, over and above that available from 2-stage springs, was necessary. The most feasible solution was an entirely new approach to existing cam action variable rate

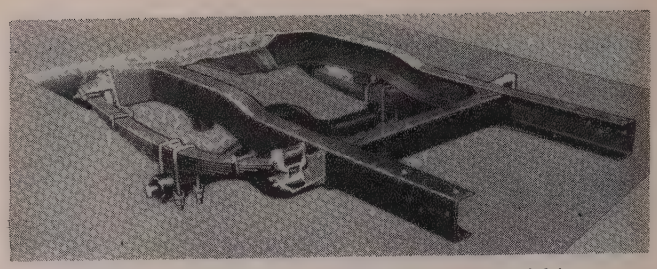


Fig. 51 — Rear suspension (1- and 1½-ton models)

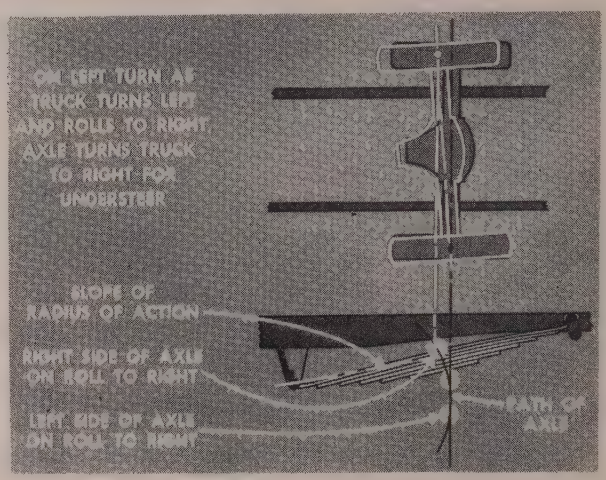

ON LEFT TURN AS TRUCK TURNS LEFT AND ROLLS TO RIGHT, AXLE TURNS TRUCK TO RIGHT FOR UNDERSTEER
SLOPE OF RADIUS OF ACTION
RIGHT SIDE OF AXLE ON ROLL TO RIGHT
LEFT SIDE OF AXLE ON ROLL TO RIGHT
PATH OF AXLE

Fig. 52 — Roll steer of rear axle

springs (Fig. 53).

These are unique in that even interleaf friction can be compensated. Since the deflection rate is entirely dependent on the cube of the usable spring leaf length, the unloaded ride can be designed to match the front springing. Under load, the spring pads can be so designed as to accomplish the same purpose.

As a further refinement, these springs are of a 2-stage design. These, through leaf lengths and the number of leaves, allowed a complete latitude of choice in rates, to complement all the advantages of the front suspension.

Roll Resistance — Variable rate springs of the cam spring hanger type, in addition to providing substantial rate variation when the vehicle is depressed, give a lifting effect on the vehicle equivalent to the vertical change in cam height, since the spring rides downward on the cam. This provides for excellent roll stability with reduced roll angle, because the side of the truck rolling downward is lifted and the side rolling upward is lowered by the cam shape (Fig. 54). This results in improved stability and driving feel.

Variable Rate Spring — As in conventional design, the new spring leaves are placed in a vertical pile, with each lower leaf progressively shorter in length. Leaf width is 3 in. and lateral spring centers are 40 in. apart. The spring piles consist of 9 – 11 leaves, depending upon the size and rated capacity of the truck. Unlike conventional rear spring assemblies, however, the new springs use no shackles, bushings, or pins at the leaf ends.

The spring hangers (Fig. 55) are riveted to the frame side rails, with the front hanger of cast malleable iron while the rear is an Armasteel casting.

Fig. 53 — Rear suspension (2- and 2½-ton models)

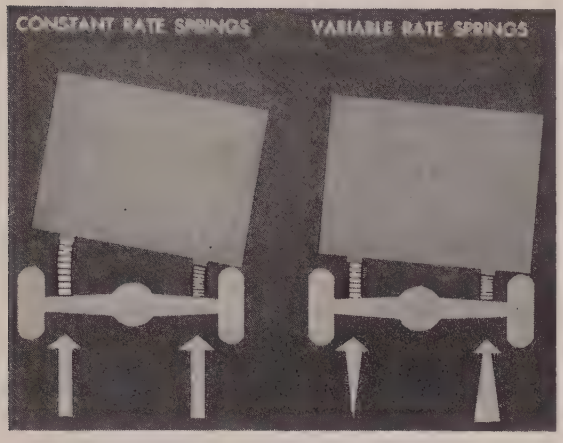


Fig. 54 — Greater resistance to roll

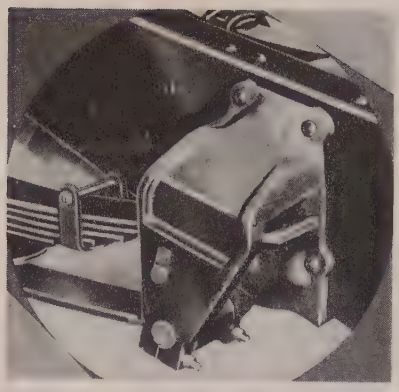


Fig. 55 — Hanger detail

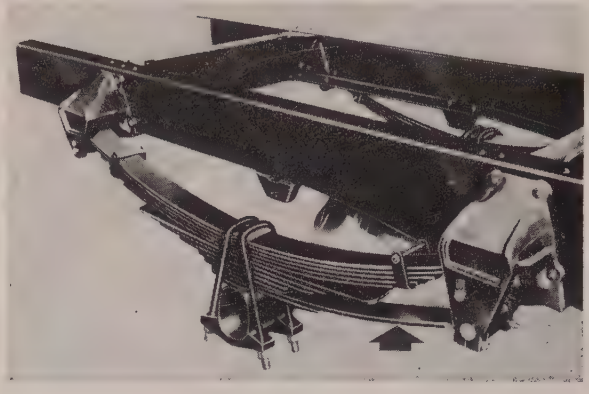


Fig. 56 — Radius rod

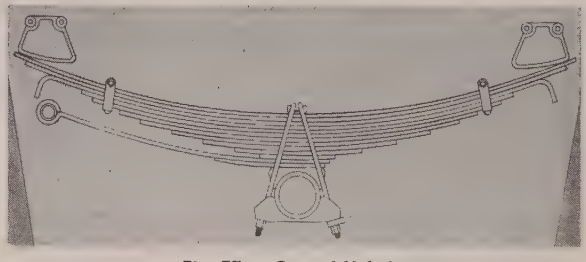


Fig. 57 — Canted U-bolts

The contoured contact pads of the rear hanger on which the springs ride are induction hardened. Both hangers incorporate rebound pins and the front component includes the pin for the radius rod.

In function, only the top leaf contacts the cams which are contoured to provide the greatest effective leaf length when the spring is at the maximum positive camber at curb load. As the spring deflects under load, its camber decreases and the spring effective length is decreased, not only of itself but because it contacts the cams at points progressively closer to each other. Since spring rate is inversely proportional to the cube of its effective length, that is the length between the cams, progressive rate is readily evident.

Hooked ends of the second or third leaf contact rebound pins in the hanger, to maintain suspension position in cases of severe rebound. Also, in the event of radius rod failure, the vehicle is completely safe because the hooked ends will engage the rebound pins. The rear pin on each side of the vehicle incorporates a rubber tube section which is used to reduce any slapping noise of an unloaded vehicle on a "washboard" road.

All leaves are shotpeened and the second stage leaves have roll-tapered ends to improve bending stress and reduce tip friction. At the bottom of the spring pile a spacer separates the last leaf from the radius rod (Fig. 56). The latter is actually a half-leaf, extending forward from the spring pile and attaching through its eye to the forward hanger. A stress plate, below the radius rod, distributes bending stresses through the radius rod leaf. The radius rod leaf and stress plate are cambered to reduce the bow and arrow effect or rebound snap by having zero rate effect at curb and a maximum rate effect at design load. This is also instrumental in controlling braking and acceleration windup and maintaining suspension alignment.

The U-bolts attaching the spring pile to the axle are canted to allow the utilization of the greatest possible active leaf length (Fig. 57), and are held in the slanted position by a short, cast spacer. To eliminate stress concentrations at their corners, and to minimize the loss of torque through the bending of the bridged surface, the bolts are arched across the top, with large corner radii. Anchor plates are of cast malleable iron and do not loosen because of the canted U-bolts and spherical nuts.

Variable rate springs lend themselves ideally to use with the independent front suspension. Differences between front and rear spring rates in heavy-duty conventional suspensions are generally quite large. Because of this variation, fore-and-aft pitch exists to an undesirable degree when the unloaded vehicle encounters bumps or rough roads.

The rate of these progressive springs has been designed to match the frequency of the front suspension with the vehicle unloaded and yet have the rate increase rapidly enough to maintain a compati-

ble frequency with the front suspension. This rate increase is consistent with the requirements for handling the prescribed heavier loads.

Roll Steering — Unlike the commercial vehicles, large trucks require less rear suspension roll steering for the cornering coefficient of large tires is far superior to smaller units. In addition, dual rear tires naturally tend to run straight (Fig. 58) because of the distance between each pair of tires running on the same road surface. Any turning of a dual rear wheel is resisted by a couple at each pair of wheels.

Large truck rear suspensions are, therefore, designed for neutral steer loaded and to be perhaps as high as 6% oversteering when unloaded.

The wheel rates obtained on the heavy-duty models are again an indication of the improvement in ride. Shown in Table 3 is a comparison of the 1959 versus the 1960 rates, with a better than 25% reduction in all cases.

A first stage spring rate versus unsprung weight relationship must be adhered to, to prevent axle hop on rough roads when shock absorbers are not used. For the first time, the suspension design is such that shock absorber control is very effective.

Tandem Suspension — Tandem axle models utilize the same basic independent front suspension as do the other heavy-duty models. Complementing their durability and ride characteristics is a new rear suspension.

It is a very durable and easily serviced unit which uses a variable rate leaf spring, equalizer beams, and is rubber bushed throughout. The torque arms are mounted to a centrally located cross-member, riveted securely to the frame flanges. The adaptation is such that no propeller shaft torsional rattles exist. This suspension is of proved durability and has been further improved through application of new and softer rear springs, as well as a change in design of the equalizer beam adapter bolts. These were previously 2½ in. in diameter and employed dished lock nuts which required 400 lb-ft of torque to install, and considerably more to remove after rusting in the field. The new design utilizes two shoulder-type castings which engage the ears of the axle brackets and the rubber bushing of the walking beam boss. They are simply held together with ¾ in. diameter drawn bolts and elastic stop nuts.

Frames — The drastically new suspension designs necessitate completely new frame configurations. The most radical departure from conventional truck frame design is perhaps represented by the assemblies used on the ½- and ¾-ton models (Fig. 59).

Side rails are contoured to provide kickups over both the front and rear suspensions. Benefits derived from the low, drop center section thus formed include a low vehicle silhouette for ease of entry and a lower c.g. for enhanced vehicle stability. Frames have varying widths, front and rear, with the wide rear section providing a stable load platform. The entire frame, including cross-members, is of all-welded construction.

Rigidity of the configuration is achieved through the cross-member construction. Behind the engine rear support, the inner halves of the frame side rails bend toward the center, where they are joined by arched plates, top and bottom, with two internally formed bulkheads, welded to form a rigid tunnel for the propeller shaft. The cross-member then continues outward and rearward, rejoining the side rails and once again forming a box-section. Two reinforcing crossbraces attach the side rails to the tunnel of the cross-member.

The frame side rail fore-and-aft areas are box-sections, approaching the rigidity of a tube.

In actual test, the torsional rigidity of this particular type of frame shows over a 1000% improvement. The new frame twists at the rate of approximately 1600 ft-lb per deg, in comparison to the 120 ft-lb per deg for the former design. Beam strength is also improved by 21%.

Substantial increases in section moduli have also been effected. The ½-ton side rail modulus is increased from 2.54 in.^{3} to 3.45 or a 35% improvement. The ¾-ton model change is from 3.37 to 3.91 in.^{3} or a 16% improvement.

The 1- and 1½-ton units retain the familiar riveted ladder-type configuration. As on the light frames, however, kickups over the front and rear wheels reduce the vehicle overall height and c.g. (Fig. 60).

The frames are built of open channel side rails, joined by rigid cross-members, and reinforced by alligator-jaw braces.

The torsional rigidity and beam strength are comparable to those of previous design, with strength improvements in the area of the front suspension. The deep, box-section front suspension cross-member, bolted to the frame side rails imparts additional strength to the entire forward frame area.

Heavy-duty 2- and 2½-ton models utilize several different riveted frame designs, each closely tailored to the wheelbase, vehicle type, and carrying capacity (Fig. 61).

Ladder-type frames are used on 121 – 157 in. wheelbase 2-ton models, the 97 – 145 in. 2½-ton units, and all school buses. A cross-member reinforced frame is provided for the longer wheelbase 2- and 2½-ton trucks, which has a torsional stiffness of 2500 ft-lb per deg.

Forward side rails in both designs are strengthened by a welded inner liner, forming a box-section member. On the lighter, shorter frames the inner linner terminates in a sturdy bridge-type cross-member to which it is attached.

The remaining intermediate cross-members are conventional flanged channels, some reinforced by alligator-jaw braces. The rear cross-members are channel sections, with gusset reinforcements.

The alternate frame construction is identical to the ladder-type in the forward and rear area. Intermediate cross-members are replaced by a rigid X-member.

The X-member arms are formed by a continuation of the inner liner of each box-section side rail. These inner rails bend toward the frame center, then parallel each other for a short distance. Upper and lower reinforcement plates are riveted to the parallel members, forming a central tunnel which is further reinforced by fore-and-aft bulkheads sealing the ends. From this strong central section the inner rails rejoin the outer side rails to form short box-sections to the rear spring front hangers.

An L-shaped reinforcement is available optionally for the 2- and 2½-ton models to increase the section modulus further.

Frames used for tandem axle units are similar

to those of other heavy-duty models. Frame reinforcements, however, are provided as standard equipment. These are partial C-channel sections up to the rear spring front hangers, continuing from there rearward as L-sections.

The program of frame tailoring to vehicle size provides frames with a number of different section moduli (Table 4). The ladder-type frames used on 2-ton units have a maximum section modulus of 11.28 in.^{3}, with a section behind the cab of 8.28 in.^{3} The maximum section represents a 19% increase over comparable 1959 components.

The X-member frames have maximum section moduli of 14.79 and 15.91 in.^{3}, respectively, for the 175- and 197-in. wheelbase models. This is an increase of 57.2 and 53.6%.

Similar increases in strength are common to the 2½-ton models. The ladder-type frame is now a maximum of 12.96 in.^{3}, or a 10% increase, while the X-member frame has a section modulus of 14.79 for the 157- and 175-in. wheelbase and 15.91 in.^{3} for the 197-in. wheelbase chassis.

Section modulus provides an indication of relative strength of a given cross-sectional area. In a program of complete redesign and re-evaluation, it may become evident that certain components have been overburdened. Such was the case with frames. Thus, some 1960 frames may be somewhat lesser in torsion and beaming, but only in comparison to the previous components. In every case, the frames have been exhaustively tested and proved more than adequate for the work they must perform.

Wheels and Tires — Any discussion of the ride, handling, and vibration characteristics of a truck must involve the tires and wheels as these components, through their direct contact with the road, are in the front line of any disturbances.

The problem of balance in rotating parts has existed since the invention of the wheel, but it has become compelling only within the past 15 – 20 years. At a time when rpm were counted in the low hundreds, comparatively large amounts of unbalance would have little effect upon the function of the machine. However, as rotating speeds become sufficient to generate significant force, better balance becomes essential.

Since centrifugal force due to unbalance increases as the square of speed, unbalance, and the forces it excites, is of great importance in the modern high-speed vehicle.

It was suspected, for some time, that many of the ride and handling troubles resulted from excessive physical unbalance and runout conditions that exist in wheel assemblies. In 1956, the GM Research Staff conducted a study of front wheel and tire assemblies and their rotating components. The front wheels only were investigated, since disturbances in the front end can have a more direct effect on the road behavior of a truck. These are transmitted to the driver through vertical vibratory displacements of the frame and body and/or as an oscillatory motion of the steering wheel. Rear wheel and tire assemblies, confined essentially to vertical motions, can transmit only the effects of vertical vibrations to the suspended body mass.

The objectives of the study were twofold: (1) to determine the amount and effects of unbalance and runout in the wheel and tire assemblies, and (2) to investigate methods of improving the balance of

Table 4 — Maximum Frame Section Moduli

Vehicle	1959 (in.^{3})	1960 (in.^{3})	Vehicle	1959 (in.^{3})	1960 (in.^{3})
½ Ton	2.54	3.39	2 Ton	9.41	11.28
	3.37	3.88			14.79
¾ Ton	3.37	3.91		10.36	15.91
1 Ton	5.70	5.09			12.96
1½ Ton	8.28	6.28	2½ Ton	11.79	14.79
	9.41	7.29			15.91

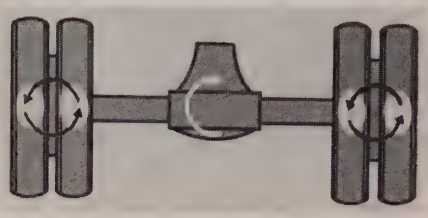


Fig. 58 — Dual rear tire yaw damping (for dual wheel axle to turn, sliding must occur at road between each tire which results in dual rear wheels wanting to go straight)

Fig. 59 — Frame (½- and ¾-ton models)

Fig. 60 — Frame (1- and 1½-ton models)

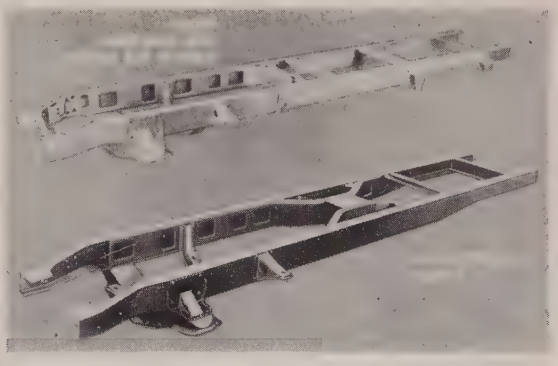


Fig. 61 — Frames (2- and 2½-ton models)

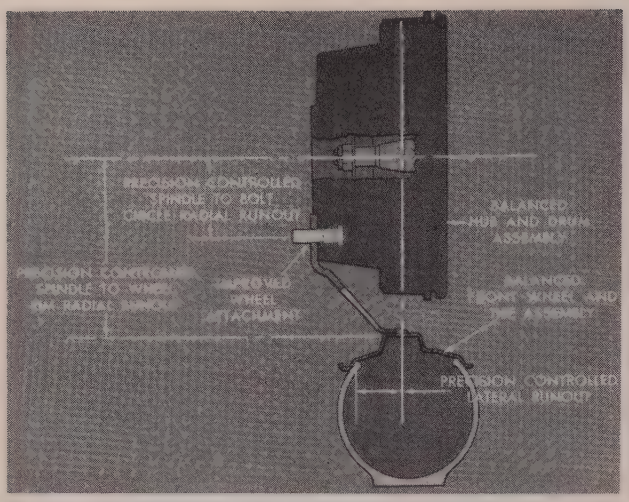

PRECISION CONTROLLED SPINDLE TO BOLT CIRCLE RADIAL RUNOUT
BALANCED HUB AND DRUM ASSEMBLY
PRECISION CONTROLLED SPINDLE TO WHEEL PILOT RADIAL RUNOUT
APPROVED WHEEL ATTACHMENT
BALANCED FRONT WHEEL AND TIRE ASSEMBLY
PRECISION CONTROLLED LATERAL RUNOUT

Fig. 62 — Manufacturing tolerance

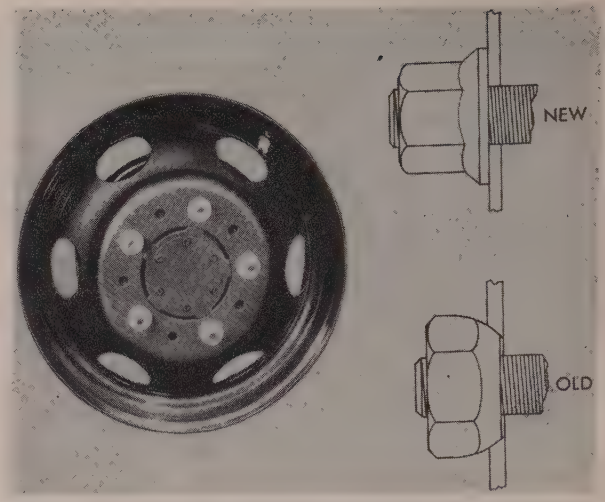


Fig. 63 — Wheel pilots

these assemblies without resorting to the expense of purchasing equipment and providing personnel for balancing components. Twenty-four sets of components were tested, in what is believed to be the first such investigation of its kind.

In this study, hubs, brake drums, wheels, and tires were examined separately and in assembly. All tests were performed on a research balancing machine, while special test setups were utilized to determine the amount of runout in the various components. These conclusions provided a basis for 1960 design.

The average physical unbalance in the hubs was found to be 8.73 oz-in., while the runout averaged 0.037 in. In 11 instances, the diameter which locates the brake drum was 0.0022 in. oversize, a condition which could account for 2.96 oz-in. of unbalance in the hub and drum assembly. In addition, the studs, which are pressed in the hub for wheel mounting purposes, were cocked causing a misalignment of 0.030 in. off-center. In a final assembly of a 160-lb wheel and tire unit, this misalignment could cause 76.8 oz-in. of unbalance.

The brake drum pilot diameter, which locates the drum on the hub, was tested for size and found to be within the accepted tolerance. However, a check of the unmachined exterior surfaces of the drum indicated an average runout of 0.070 in. Physical unbalance in the 24 test pieces averaged 42.66 oz-in.

The wheels were mounted on a special arbor and checked at the bead seats where the average runout was determined to be 0.048. in. The average unbalance reading of all wheels tested was 20.22 oz-in.

Tires were found to be the major cause of total unbalance in the front end assembly. Used in the test were 10 – 22. 5 – 10 ply rating tires. The wall thicknesses were found to vary 0.096 in. on the average, while runout averaged 0.055 in. on the outside diameter and 0.083 in. on the face of the casing. Tire out-of-balance ranged from 126 to 260 oz-in.

The second objective of the study, namely that of developing a method of improving the balance of the wheel assembly, without resorting to a balancing operation, proved unsuccessful. Of the several methods tested, each proved impractical.

A close evaluation of the Research Staff report led to the realization that the only feasible solution to the unbalance and runout conditions lay in improved designs and new, higher quality requirements for the front end rotating components, with special emphasis placed on the tire assembly. In addition, in order to keep total unbalance within newly specified limits, an assembly plant balancing operation for all front hub, brake drum, wheel, and tire assemblies was necessary.

Radial Runout — The first step was to improve front wheel concentricity, or radial runout, by properly centering the wheel on the hub (Fig. 62). To accomplish this, commercial vehicle hub assemblies were provided with new manufacturing tolerances which control the bolt circle-to-spindle runout to 0.010 in., and a wheel rim bead seat-to-bolt circle runout of 0.060 in.

On the larger medium and heavy-duty models, where a greater mass of weight and larger radii amplify the out of balance condition in the front wheel assembly, disc wheel hubs, and cast spoke wheels are redesigned to incorporate a wheel or rim pilot (Fig. 63). On the disc wheel hubs, five equally spaced shoulder lugs pilot the wheel inside diameter. To minimize eccentricity, a tolerance of 0.005 in. is specified for hub pilot-to-spindle runout. The design and spacing of the lugs are such that a tool can be inserted between the wheel inside diameter and hub for easy wheel removal, even after rust buildup. The rim pilot on cast spoke wheels consists of a machined seat at the outer extremity of each spoke. Rim pilot-to-spindle runout is limited to 0.003 in.

Another feature which provides a significant contribution to wheel concentricity is the new wheel attachment on the medium-duty models. The tapered nut, formerly utilized on these vehicles, is discontinued and replaced by a washer-based nut attachment. It was found that the tapered nut could not be depended on to perform its dual task of securing and piloting the wheel on the hub. Minor variations in the bolt circle, or in the concentricity of the attaching nut itself, affected the true running performance of the wheel. Fig. 64 illustrates three conditions which cause wheel eccentricity in a tapered nut attachment. Cocked studs, runout in the bolt circle, or an improperly machined attaching nut caused the wheel to shift

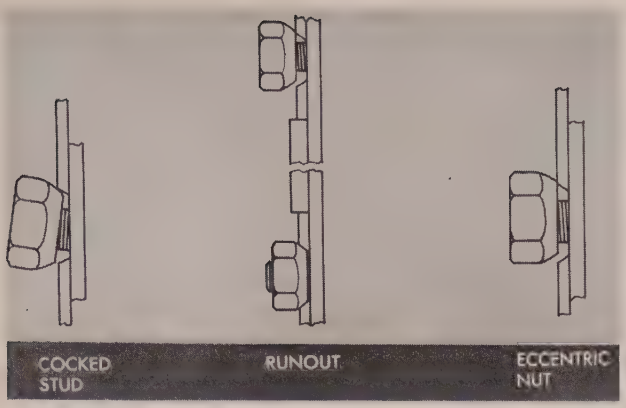


Fig. 64 — Tapered nut attachments

as much as 0.060 in. on the hub, resulting in a severe radial runout.

In the new attachment, the wheel is properly centered on the hub by piloting lugs, and the washer-based nut performs the single function of firmly securing the wheel to the hub-and-drum assembly. The attaching nut provides a wide contact area with the wheel, thus reducing the stresses at the stud holes. The new wheel assemblies are not interchangeable with those used with the tapered-nut attachment. However, because of many advantages afforded by this wheel-mounting system, a rapid industry-wide acceptance of the washer-based nut attachment is anticipated.

On heavy-duty models equipped with disc wheels, the tapered-nut attachment is used in conjunction with the pilot lug feature. However, the lugs, not the nut, are relied on to pilot the wheels.

Lateral Runout — To control front wheel wobble, or lateral runout, a 0.005 in. tolerance is specified for the hub mounting surface-to-spindle dimension, while the wheel rim bead-to-mounting surface runout is limited to 0.060 in. on the commercial models and 0.090 in. on the heavier vehicles.

All front hub-and-drum assemblies are statically balanced to assure true running performance of the front end rotating components. Previously uncontrolled, out-of-balance in the front hub-and-drum assemblies is now limited to 6 oz-in. on the commercial models and 10 oz-in. on the medium- and heavy-duty vehicles.

Tires — Tests indicated that a vast improvement in the uniformity and balance of tires and wheels was needed, and to this end, new and more exacting tolerances were specified to attain a uniformity of construction beyond anything achieved before.

To keep total runout and balance within the specified limits, radial and lateral tolerances for wheels were established. Of the total, 90% of wheels are to be within 0.050 in. and the remaining 10% to be within 0.060 in. It was also requested that an out-of-balance limit of 30 oz-in. be maintained on all wheel assemblies. Improvements in wheel assembly and rim processing methods were developed which indicated that the wheels could successfully be manufactured to the specified tolerances.

Since the tire casing was found to be, by far, the greatest contributor to the total unbalance of the front end assembly, it would appear essential that a definite improvement be made in the balance and runout condition of this component. Although tires appear to be quite simple, the inherent quality of flexibility makes these structures highly complicated and somewhat unpredictable. It is the contention, however, that not enough attention has been paid to the truck tires, with respect to their ride and handling characteristics.

A case in point is the troublesome problem usually referred to as "tire thump." This is a vibratory condition definitely excited by the tires. Investigations have established that there is a correlation between thump and tire deflection uniformity. During the last 5 years, much has been done in the way of development of new processes and more precise equipment to minimize tire thump in the passenger-type tire sizes. However, these advantages were not instituted for truck tires, until we requested this same uniformity of construction in the truck tire assemblies.

The tire builders, accepting the main responsibility for maximum tire uniformity, instituted a program to build truck and bus tires to an improved balance construction.

Preventive measures alone, however, were not expected to achieve the condition of balance required in the front end assembly and a final corrective operation, factory balancing of all front wheel-and-tire assemblies, was introduced to reduce unbalance to the desired limits.

Balancing Procedure — There are two general types of wheel-and-tire balancing machines available, the static and dynamic. The first measures and locates weight causing unbalance in a single plane, the other in two planes. Both the machines are practical compromises as the number of planes in a mass is neither one nor two, but infinite. However, considering the objective of any balancing operation, the perfectly smooth operation of the part, it is valid to consider most revolving parts as operating in either one or two planes. As a matter of fact, if the areas of correction do not lie in distinctly different and comparatively widely separated planes, it is impossible to accomplish anything but single plane balance.

A factory balancing operation requires accuracy. However, speed and economy are also significant. Our assembly plants are equipped with the latest design Micro-Poise machines. These are static precision balancing units, utilizing a fixed base and pivoted platform which is free to tilt in any direction without mechanical friction.

When the operating handle is released, a condition of unbalance in the wheel and tire assembly will cause the pivoted table and workpiece to tilt in direct proportion to the amount of unbalance. An optically precise universal level mounted on the platform amplifies the tilt, indicates its direction and measures it in terms of in-oz. A correction to the out-of-balance condition is made by adding a weight of the indicated amount 180 deg away from the direction of the unbalanced mass.

In the final analysis, the closer manufacturing tolerances and balanced wheel-and-tire assemblies are but an important part of the overall redesign, contributing to improvement in vehicle ride and handling.

Durability

Any completely new product naturally involves a question of durability from the consumer's stand-

point. The monumental investment in time, effort, manufacturing, engineering, and tooling costs, however, demands a complete assurance of durability. There is no room for error.

It would be far more expedient for a large corporation to continue existing design than to risk loss of prestige and reputation with a product not proved in all conceivable respects.

The development of the 1960 Chevrolet trucks was the result of thousands of tests in the laboratory and hundreds of thousands of road miles involving both the proving ground and actual road trips.

As stated previously, a set of safety factors and baselines was established which the new designs had to meet. Every component, every assembly, every unit had to complete a schedule of tests, on a comparison basis with previous design. The durability of the latter was a known factor which could be translated into theoretical vehicle life.

A complete description of the entire testing program would entail entire volumes. Following, however, are representative samples of laboratory and proving ground tests, the results of which are typical for the entire program.

Laboratory Tests — To determine the loading for brake reaction tests, a truck of each series was loaded to 112½% of the rated gvw, with the appropriate load on the front suspension. Various areas were instrumented and a number of emergency brake stops were made. From this information, the laboratory test loading was accurately determined.

An interesting sidelight is that during these stops the coefficient of friction between tire and pavement approaches unity. Also, on all models above the ½-ton category, the load on the front suspension doubles.

All seven suspensions were laboratory tested in this manner and proved superior to 1959. For example, the 1953 ¾-ton suspension, which is very acceptable in service, failed the front springs in 76,000 cycles. The 1960 suspension was in good condition after 300 thousand cycles and survived an additional 100 thousand cycles with a 25% increase in load.

Fig. 65 shows the sand and heat test to which the independent front suspensions are subjected. No comparable tests were conducted on 1959 components since this is basically a test of wear on the bearings of the pivoting members. The suspensions are tested a minimum of 2½ million cycles, at full stroke, bump rubber to bump rubber. This is considered an equivalent of hundreds of thousands of miles under actual operating conditions.

The torsion bar cycling test requires one million cycles, without failure, before a bar is released for production. Many tests, however, totaled more than twice that amount. In comparison, the 1959 components averaged a total of 170,000 cycles before some failure occurred.

Coil springs, as used on the rear suspensions of the ½- and ¾-ton models must also display their durability. A total of 100,000 cycles, considerably in excess of 1959 requirements, is necessary before the springs can be released as production parts. Tests are generally conducted to 250,000 cycles, as an additional guarantee of reliability.

The variable rate spring was also tested. Similar trials were never conducted on previous design so no direct comparison can be made. However, an educated guess, utilizing data gathered at the proving ground, places the cycling durability of the 1960 component at four times that of previous spring design.

An impressive test of rear suspension ruggedness is labeled the Rock and Roll. The chassis is loaded with 20,000 lb of iron and rocked back and forth by the cycling machine. A total of 18,000 cycles was absorbed by the 1959 components before spring failure. The 1960 chassis withstood a total of 66,000 stresses before failure. This, however, was only a

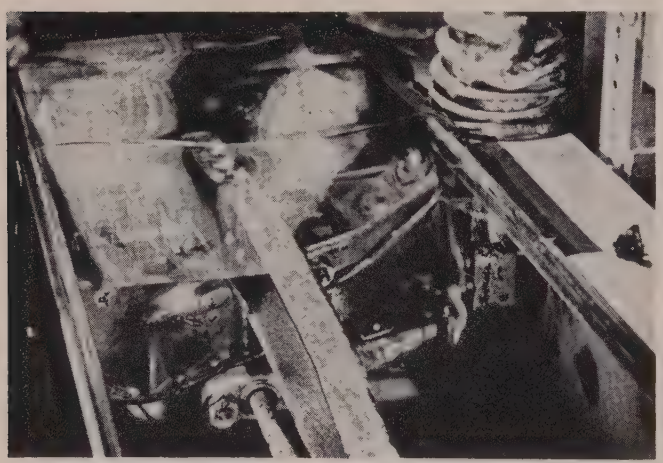


Fig. 65 — Sand and heat test

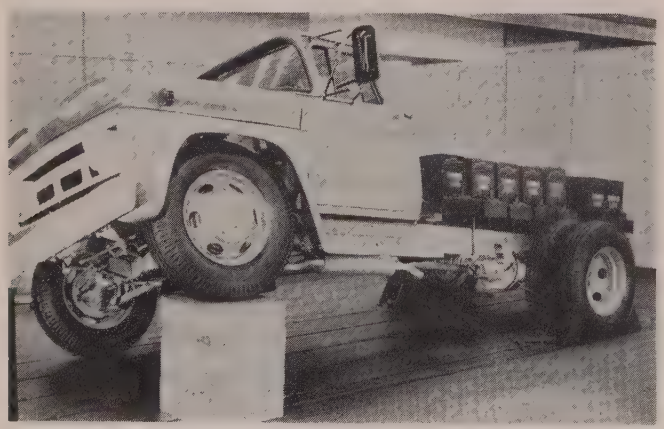


Fig. 66 — Chassis deflection test

Fig. 67 — Railroad ties test, ½-ton vehicle

loosening of the rivets attaching the spring hangers. This deficiency has been since strengthened by a hanger redesign.

To establish the effect of chassis deflection on the cab and sheet metal, (Fig. 66) one front wheel was raised a total of 35 in. off the floor. All doors, and the hood maintained their relationship to adjacent components and could be opened and closed without additional effort.

Proving Ground Tests — The testing program does not stop in the laboratory. Additional exhaustive tests are continued at the proving ground, over every conceivable type of operating terrain.

Typical is the Belgian Block schedule. The torturous course is built of cobblestones, and incorporates additional vehicle stressing ditches and bumps. A normal test schedule calls for 3500 vehicle miles. Previous models, however, were capable of completing only one-third to one-half of the schedule without suspension failures. It is typical of the new models to complete this entire test without failures, and some vehicles have gone around two and three circuits. Additionally, the overloads are more closely adhered to, and it is possible to run the vehicles at a higher safe speed. Considerably less maintenance is required to keep the vehicles running, and at the end of the test the trucks are in a much better condition.

Special tests were also conducted to test operational qualities in off-the-road operations. In this particular experiment, the vehicle drove into the sea of mud under its own power, and when it lost traction, a tractor pushed and pulled it through the remaining area. Once out of it, the truck, undamaged, continued under its own power.

Shown in Fig. 67 is a ½-ton vehicle negotiating the railroad ties test. This is a super-duty experiment, never before conducted, which was initiated to study further the effects of rough roads on the suspension systems, since the Belgian Blocks resulted in no failures. Fig. 68 shows a larger model negotiating this same course.

Another addition to the Belgian Block schedule is the cross ditches which the vehicles negotiate once every time around the circuit. Diagonally opposite wheels alternately enter the ditches in a test of the chassis durability.

The preceding are but a limited representative sample of the truck testing schedules conducted at the laboratory and at the proving ground. In addition, many over-the-road trips are conducted to test vehicles under actual operating conditions. The fact remains, that the durability of the 1960 models is far and above that of the 1959 units.

"Putting a Number" on Actual Improvements

Much has been said about ride improvements in the theoretical sense. The hypothetical advantages of independent front suspension and softer, frictionless torsion bars, the 2-link rear suspensions and coil springs, the variable rate springs, all indicate tremendous improvements in comfort.

There has been a great deal of discussion on lower rates, lower frequencies, and higher amplitudes and their relation to ride and comfort. It became evident in the development program that some method of rating the actual improvements had to be adopted to prove that the objective was obtained.

Instrumentation — To this end a system of instrumentation was devised which permitted ride evaluation on a comparative basis. The system (Fig. 69) included appropriate accelerometers as the primary source of signals. In an effort to make the data acquisition system as light and portable as possible, the accelerometers were powered and demodulated by a transistorized 5-kc carrier system. The outputs of the accelerometer demodulators were each used to modulate one of the five voltage-controlled oscillators whose frequency-modulated outputs were recorded on one channel of a 2-channel tape recorder.

A separate vehicle carried the recording equipment which was connected to the sending units in the test vehicle by a cable.

The accelerometers were strapped to the driver's chest and hip. Comparable 1959 and 1960 vehicles were driven, at the same speed and by the same driver, over the same stretches of chosen rough and smooth roads. The severity and frequency of accelerations were recorded on one channel of the tape recorder. Simultaneously, the outputs of two

Fig. 68 — Railroad ties test

Fig. 69 — Instrumentation for ride evaluation

fixed-frequency oscillators, used for tape speed compensation were recorded, along with voice commentary on the other channel.

Data thus obtained was later played back through FM discriminators and reduced, utilizing an analog computer. The reduced data was then recorded in tabulated form and curves plotted of acceleration levels, versus the number of occurrences.

The taped data can also be fed into an oscillograph (Fig. 70) to provide a visual representation of the recorded ride.

Many different methods of data evaluation were considered. Peak values of acceleration would not be representative of the general vibrational characteristics of a vehicle since peak values would base the evaluation upon a single value, such as a large bump. Average values of acceleration are also unacceptable, since one vehicle which vibrated continuously at a rather low level might have the same average vibration as another vehicle which vibrated only occasionally, but at a much higher amplitude.

It was, therefore, decided to reduce the data to a statistical distribution presentation which would consist of a plot of levels of acceleration, versus the number of times during the test that the acceleration amplitude exceeded the corresponding level.

For the sake of brevity, this presentation is limited to the evaluation of only four vehicles, since the results are representative of the entire study. Tested were 1959 and 1960 ½- and 2-ton units. The tests were conducted at the proving ground, over the same stretches of both the smooth and rough road, by the same driver, at the same speed.

Acceleration Curves — Fig. 71 is a curve plotting vertical motions, with the accelerometer located at the hip of the driver and the vehicles traveling over the rough road. On this and the following curves, the accelerations, in *g* units, are plotted along the ordinate, while the total occurrences of a particular *g* value are plotted along the abscissa. The areas below the two curves were calculated to gain some idea of the improvement and show a decrease of 53% in the area below the 1960 curve as compared to the 1959.

Fig. 72 shows the results of this same test on larger 2-ton vehicles. Again, there is a considerable reduction of 58% below the curve depicting the 1960 ride as compared to the 1959.

Tests on smooth roads (Fig. 73) follow the same pattern. Although there is a reduction in the number and severity of vertical accelerations for both vehicles, the 1960 model shows an improvement of more than 78%.

The same is true for the larger models (Fig. 74). Here an improvement of 52% is shown.

In the areas of fore-and-aft motions, where the accelerometer was strapped to the driver's chest, similar reductions in relative motion are common.

Fig. 75 depicts the curves for a 1959 and 1960 ½-ton unit, running over the rough road. The curves here indicate a 68% improvement. Heavier models (Fig. 76) follow this same pattern and show a 42% reduction for the 1960 model.

The vehicles were then conducted over a smooth road, with the same results (Fig. 77). The ½-ton models for 1960 indicated a 70% improvement, while the 2-ton units are improved by 35% (Fig. 78).

Critique — It must be noted, that unlike the studies presented at the beginning of this paper, no effort was made to define comfort.

Several of the published studies considered seated subjects experiencing vibration only in the vertical direction. Any study that does not include vibrations in all directions is inconclusive, as it implies that excitation at any level in the neglected direction is of no consequence.

Many studies are restricted to continuous repetitive excitations and deal with average amplitude of

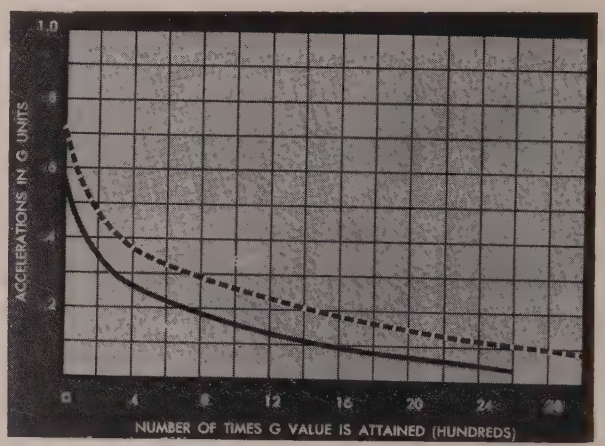


Fig. 71 — Vertical motions at hip (½-ton models on rough road)

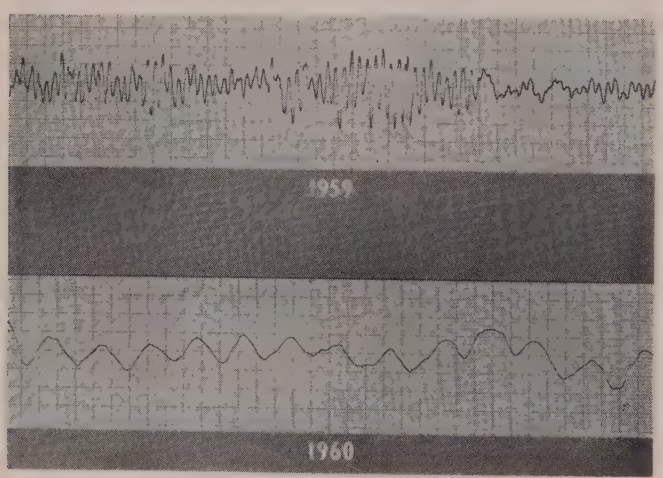


Fig. 70 — Oscillograph data of ride

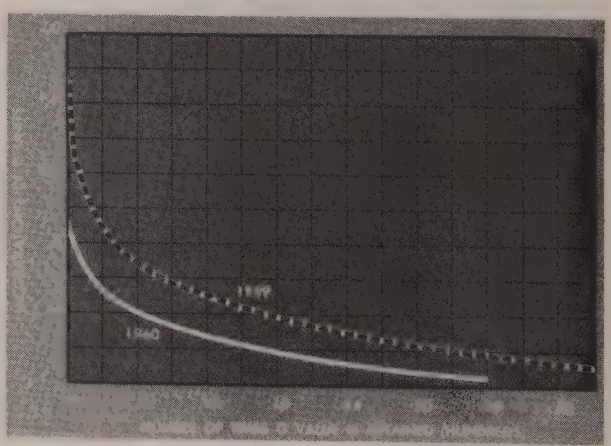


Fig. 72 — Vertical motions at hip (2-ton models on rough road)

displacement, acceleration, velocity, or rate of change of acceleration. If the study considers single frequency sinusoidal vibrations, then it matters not at any one frequency whether the vibration is considered in terms of acceleration, rate of change of acceleration, displacement, or the velocity since these are all dependent qualities, related by various powers of the circular frequency of vibration.

If the motion considered is of the nonsinusoidal variety, a square wave or random function for an example, it is customary to discuss the response of the subject in terms of his response to each of the sinusoidal harmonic frequencies which comprise the motion. At this point, the study may just as well be restricted to sinusoidal excitation because invariably the phase relation of the various harmonic components is lost. Many such methods make no allowance for the presence of two separate frequencies occurring simultaneously.

Many published correlation methods are based on root mean square, or average levels of vibration and thereby imply that there is no difference between a continuous vibration at small amplitude and a discontinuous occasional vibration at relatively much larger amplitude but having the same average or rms level. For this reason, methods of correlating physiological responses to objective vibration data are inadequate.

The data compiled by Chevrolet has no physiological significance but is objectively correct and in units and form easily understood. The only conclusion drawn is that a vehicle which experiences

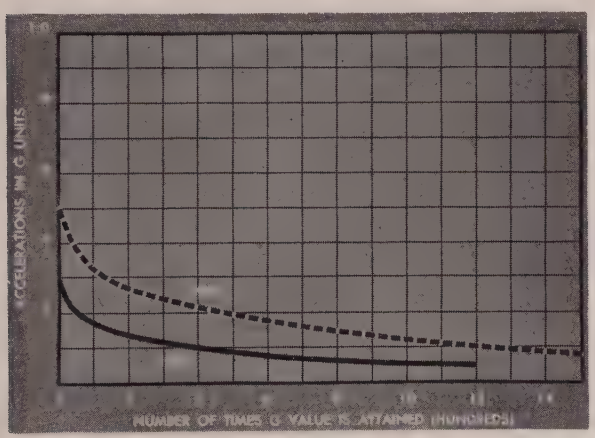


Fig. 73 — Vertical motions at hip (½-ton models on smooth road)

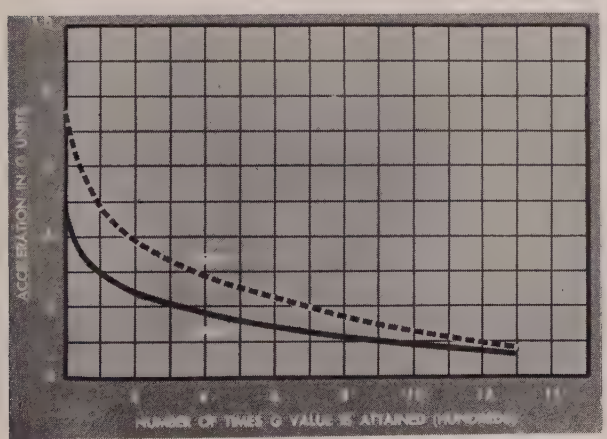


Fig. 74 — Vertical motions at hip (2-ton models on smooth road)

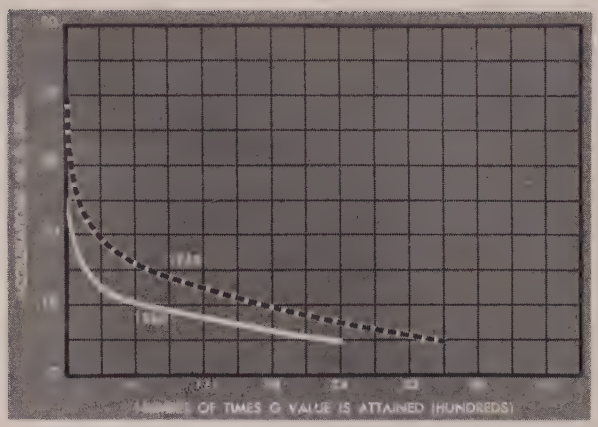


Fig. 75 — Fore and aft motions at chest (½-ton models on rough road)

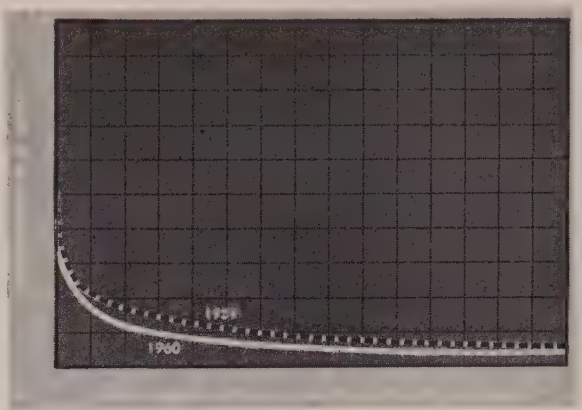


Fig. 76 — Fore and aft motions at chest (2-ton models on rough road)

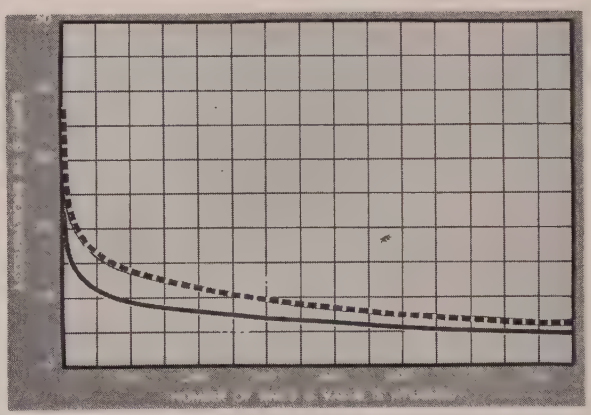


Fig. 77 — Fore and aft motions at chest (½-ton models on smooth road)

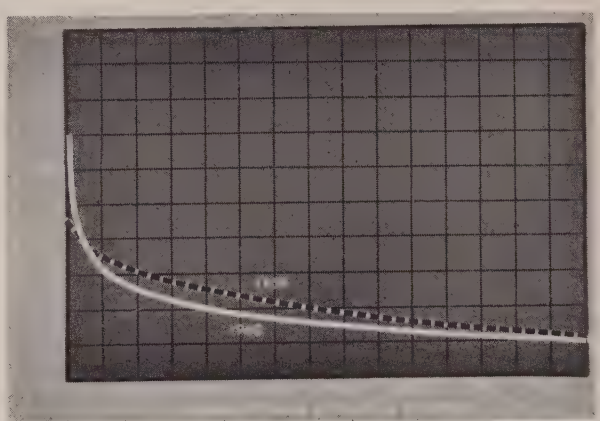


Fig. 78 — Fore and aft motions at chest (2-ton models on smooth road)

rather large accelerations in any direction, many more times than another vehicle, will quite probably be less pleasant to ride in. Tests prove the accomplishment of this objective, since the new design subjects the vehicles, driver, and cargo to fewer accelerations and of smaller intensity.

These efforts to develop the theory of obtaining a better ride with improved handling clearly indicate that this approach was practical. The actual vehicles reacted according to prediction.

It was also demonstrated that such vehicles are more durable, and provide greater comfort for both driver and cargo.

DISCUSSION

Truck Behavior Different When Used as Tractor

— E. B. Ogden
Consolidated Freightways, Inc.

THIS PAPER is so all inclusive, that it might well be used for a maintenance manual, or even for a guide in manufacturing the vehicle. Certainly all decisions made in the production of this vehicle are explained and justified.

The author apparently deals only with full-truck loaded and empty, since no mention is made of the same vehicle and its behavior when it's used as a tractor, which it certainly will be in at least 2½-ton class.

In Fig. 6, the author explains in detail the flywheel effect of the spring mass. This, in my opinion, presents a completely different picture if the vehicle is used for a tractor, rather than a truck. Many years ago we changed from a full-truck, full-trailer operation to a semi-trailer operation. In the process we removed bodies from many full trucks without changing the wheelbase of the truck, and made tractors out of them. As a truck, they were above reproach, insofar as ride was concerned; however, when they were made into tractors, the ride was nothing to be proud of.

From this experience and from more recent experience, I again state that it is an entirely different condition when the vehicle is used for a full truck with the load imposed upon the frame and solidly held in place, as against it being used as a tractor and the load imposed upon the the frame in a rather flexible manner.

The author refers to the torsion bar as the spring decided upon for all three models. Possibly had they not had the largest model to think about (the 2½-ton size), they might not have gone to torsional suspension, since coil springs might well have done the job. Coil springs could have been installed, as they have been in the past, and been within reasonable size. Frankly, as far as I know, there is little difference in benefits whether torsion bar or coil spring is used, since either method is friction free.

In regard to the rear suspension, I am somewhat disappointed, as seemingly nothing new has been added. The smaller models are using coil springs, which, as we all know, is quite satisfactory, providing you can stand the deflection. In the larger models, we come to the multi-leaf spring in which I fail to see anything new, since we have known of the variable rate leaf spring for many years. We will all admit that the variable rate leaf spring is better than the constant rate. However, I believe we must agree that to reduce the frequency in any spring, the deflection must be increased. For variable rate springs to be effective, they must have extra deflection built into them. Many times this extra deflection is undesirable from the operators' viewpoint, since the deflection is easily measured in the reduction in cubic capacity of the body.

What I'm really saying is that to receive a good ride from variable rate rear springs, from curb load to maximum gross load, there is usually a static deflection of about 4½ in. One must then add about the same amount for dynamic deflection, which is a total deflection of about 9 in. This represents the amount of clearance necessary over the rear wheels to prevent the van from rubbing the rear wheels. This, usually, is entirely more than the average operator can tolerate. I had hoped that torsion bar rear suspension was going to be used with perhaps an automatic height control, such as wes used by one of the passenger-car manufacturers several years ago. This would reduce the necessary deflection by about one-half, as the static deflection could be adjusted before the vehicle was put in motion.

The tandem axle suspension seems to be of a conventional design and if there is any action at all in the rear springs under light load, it would be a great surprise to me. Frankly, I fear there is but little action in the rear springs in this type of suspension at any time. Luckily, in this design the walking beam effect smooths out about half of the bumps.

We agree completely with the author, when he states that he would like to see closer control of wheel runout and tire balance. All operators would like to see closer control in this field. It is most heartening to see one of the leading manufacturers telling the tire companies the same story that operators have been telling them for some time.

The durability portion of this paper is most interesting, but I fear more from a merchandising, rather than an engineering viewpoint. I think it is most heartening that so much test work has actually gone into the vehicle, insofar as durability is concerned. However, I am at a little loss as to how this directly affects the ride and handling characteristics of the vehicle.

In my opinion, a most interesting part of the paper is the last section, "Putting a Number on Actual Improvements." For the first time we are getting down to the much needed action — that of measuring in a competitive manner the improvement made in the same vehicle from year to year or in competitive vehicles from year to year. Further, I think it is most encouraging that we are recognizing that there are two motions that must be measured and at two places on the operator's body. The set of motion curves in relation to the *g* units versus frequency are the first set of such curves I've ever seen. I feel that as long as the industry continues each year to improve the ride over previous models, we at some future day will reach a stage that is most acceptable.

Control of Ride Quality Affected by Vehicle Weight

— Lewis C. Kibbee
American Trucking Association, Inc.

WE in the trucking industry have certainly needed some real improvements in truck ride, and from this paper I gather that these improvements have been made. Naturally, since I have not had an opportunity to ride in the truck itself, I can only hypothesize what has been done from the description given. If this ride improvement has also incorporated design features in the truck which will improve vehicle durability, its success is assured.

Light-Duty Models

In this field I believe there is the least need for improved ride. My experience with the 1959 model indicated that truck ride for these lighter models (½-ton pickups and panels) was acceptable in highway operation. One shortcoming of the 1959 models had been the tendency of the vehicle to hop when the vehicle was used in off-road conditions or in deep snow, sand, or mud. This seemed to be due to a wrap-up of the rear springs as torque was applied to the rear axle. This hop became so severe in some conditions that the vehicle could not negotiate at all, short of severe damage to the drive line components. With the 1960 model and its trailing arms holding the axle stationary and receiving the torque reaction, this shortcoming apparently has been completely overcome in the light models using the coil springs. This is not the case for the 1-ton truck where long leaf springs are still used

and no torque arms are used. I believe this will result in the same old axle hop in the 1-ton model.

In this country, as well as in Europe, the independent front suspension on vehicles with axle weights approximating those of the ½-ton and 1-ton pickups have been in use for some years, and many of them using torsion bars for springs. It would appear from a showroom look at the "hardware" under the new Chevrolet truck that the independent suspension used in the 1960 models is certainly strong enough and perhaps even a bit overdesigned for this size. If this is the case, probably long life will result with the attendant acceptance by the owners. I myself have some reservations as to the wearing capabilities and the ability of the caster and camber adjustments to stay in adjustment without constant "tinkering." Perhaps I will be proved wrong in this area, but this type of adjustment in truck practice seems to be something that will have to prove itself.

With the use of low-rate coil springs in the rear of these vehicles, I would have additional reservations on the change in body bed height between the loaded and the empty vehicle. If the ride is all that it is said to be and the rate is in the areas mentioned in the paper, I cannot help but wonder how the vehicle will be prevented from appearing to be "tail down" in the loaded condition or "tail up" in the empty condition.

The new torsionally rigid frame incorporated in these light models should add considerably to the life of bodies applied to this series of vehicles. There should be little racking of such bodies and longer life will doubtless result, as well as no squeaks and rattles.

The method of placing the steering gear in the chassis appears quite ingenious and it is notable that the steering linkage has all been installed with an eye to protecting it from mechanical damage. All of the linkage to the front wheels from the steering gear is well up above the cross-members and out of the way of ruts, rocks, and such things that might normally damage a lower mounted arrangement. The heavy cross-member also is low enough to protect the engine oil pan from similar damage. It has also been pointed out that with this arrangement of the steering gear, a tighter steering angle can be achieved inasmuch as all of the steering linkage has been moved out of the wheel housings to get away from any restrictions. This is a forward step and will achieve wide acceptance. The one reservation I have with regard to the steering mechanism is the many lubrication points, adding to maintenance complexity. I have never experienced a truly "greased-for-life bushing" as is mentioned, but if this has really been achieved, perhaps we have made another step forward. The concept of being able to clamp the steering gear to the cab and allow it to move relative to the frame should be an improvement as far as the driver is concerned. Previously, where the steering gear was mounted on the chassis and the cab moved relative to the steering gear, objectionable wheel shake and vibration was encountered. I assume this new linkage in the 1960 models this shortcoming has been eliminated.

One question I would like to ask concerning the new 1960 models is, why have you stopped short of full independent suspension for all axles, front and rear, when you have swing axles for the front in the new model? Wouldn't full independent suspension add substantially to the favorable ride characteristics?

One other factor which would cause some concern as a potential buyer of this light model truck is the serious underbody restriction caused by the trailing arms for the rear axle. These arms are so close to the road that in ruts, sand, and deep snow they would effectively block the forward progress of the vehicle. One mitigating factor is that these trailing arms come together and are narrower at the front of the vehicle than at the rear, which will give the driver the possibility of being able to back up once he gets stuck from this restriction. Additionally, the exhaust system of the vehicle is strung haphazardly underneath the chassis and will be damaged severely in any rough country use. It is quite evident that a re-examination of the exhaust system placement is in order for the next model in 1961 in order to get the exhaust system up under the frame for protection.

An area in which I was particularly impressed by Mr. Flynn's paper is the length now being used by Chevrolet to have wheels in balance. I sincerely wish that other truck manufacturers as well as passenger-car manufacturers would take half as much pains towards eleminating wheel fight and other resultant steering ills by doing this much work on wheel balance.

Finally, with regard to the lighter models, I personally regret the trend to a low silhouette which makes the vehicle hard to get in and out of for a man working with it daily on the street. Regardless of where the floor starts, a man usually is standing on the ground and when he has to stoop over to begin to get into the truck, this is extremely tiring at the end of a day's work. The previous models had been fairly good in this respect, and I feel the lower silhouette is not a step in the right direction. Additionally, this takes away valuable space which is necessary in panel models as the dimension from the top of the floor to the bottom of the ceiling is quite important in this respect.

Heavy-Duty Models

In the heavier vehicles, it will be noted that the ground clearance is not appreciably affected by the independent front suspension in that the new heavy cross member is approximately at the same height as had been the old I-beam front axle. This cross-member likewise protects the engine crankcase and the whole aspect is a clean job, well engineered.

Due to the complex frame in the longer wheelbase vehicles, adaptions and frame extensions will be very difficult in those wheelbase lengths. I believe this will cause the 121- and 175-in. wheelbases to be popular for remodellers, as they have the conventional ladder-type frames which are easier to re-work in a body shop.

Certainly, the heavier models in this new line are an area where improved ride was really needed and the new independent suspension should prove to be a big help. At the same time we have to consider that maintenance costs must be competitive with the previous low upkeep costs experienced by the time-proved leaf-spring. If a better ride as well as reduced costs can be achieved, the unit probably will be successful.

The new type of wheel piloting on the wheel studs is interesting and this concept has been one of the several choices available for disc wheels for some years. The pros and cons of the merits of hub pilot, stud pilot, or wheel nut pilot are well-known and it seems to be a case of "pay your money and take your choice." The interest in this new development, from my point of view, is the departure by Chevrolet from its long-established practice, now introducing a completely new wheel for fleet operators.

Looking at the frame of the heavy-duty models, it is interesting to see a combination of a torsionally stiffer frame in the long wheelbase models which incorporate a cross-member and box-section siderails, coupled with the apparent ability to absorb extreme racking as shown in Fig. 66. This has been a dilemma in chassis design for years, and all I can say is that I hope that one of these factors do s not work against the other.

Any evaluation of economy of the new vehicle cannot be completely made on the basis of previous models. The general reaction from a driver will be to drive a vehicle that rides better at increased speeds until the seat-of-pants reaction is the same as the previous model. This is another way of saying that the new model will have to absorb more road shock than its 1959 counterpart and still be within the cost rane that the owner of the vehicle will accept. If it can decrease travel time over the same course and yet not have a subsequent increase in maintenance cost, the success of the design is proved.

THIS REPORT deals with the major parameters of a seal application which affect its efficiency and life, as determined by controlled laboratory testing in GM Research Laboratories.*

A. Shaft
   1. Surface Roughness
   2. Machining Lead

B. Assembly

C. Seal
   1. Seal Diameter Control
      Trim Interference
      Spring Rate
   2. Seal Lip Pressure
      Trim Interference
      Spring Rate
      Rubber Hardness
      Eccentricity
   3. Seal Eccentricity
      Mold Register
      Assembly
      Trim

Seal Testing to establish Can

Robert L. Dega
and James D. Symons
Research Laboratories, General Motors Corp.

THE GM Research Laboratories undertook a program in 1957 to investigate the causes for seal leakage and to determine basic design criteria which would enable the industry to apply seals without long and expensive development periods. Contrary to generally accepted theory, it has been established by research that there are three basic factors of equal importance which cause leakers: shaft, assembly, and seal.

A lip seal should no longer be considered as a piece of flexible material used to plug a hole around a rotating shaft, but rather as a precision-built assembly — one that requires surface finishes, fits, and lip pressures necessary to establish a lubricant film between the mating surfaces. This is not unlike a precision journal bearing, and it is evident that present uncontrolled factors, to be described, which disrupt the lubricant film are the direct cause for the majority of *shaft* and *seal* leaks. For instance, a shaft that is too smooth and fitted with a tight seal cannot sustain a lubricant film and the resulting material contact produces heat decomposition and high wear. Conversely, too rough a shaft will allow the roughness peaks to penetrate the relatively thin film, again resulting in abrasion of the seal lip.

The mechanism of sealing is not yet clearly understood but it is apparent that all of the factors that make up surface finish, fits, and pressures must be more rigidly controlled in production to assure a "controlled leak" in operation. Years of experience and research in the journal bearing industry have shown that reliable operation and long life have been realized primarily from precision fittings, as opposed to the theory that the mating parts should be assembled tight and operated to seek their own running clearance. The latter thinking still prevails in the manufacture and application of lip-type seals, resulting in numerous field leakage problems.

Shaft

If we realize that a seal operates on the same principle as a close-fitting journal bearing, attention must be given to the shaft characteristics, since both the bearing and the journal finishes are important for successful operation. In other words, some type of oil film must be sustained between the sealing surfaces. Test data indicate that this film is on the order of 5–15 microin. depending on seal lip loading.

The effects of shaft surface finish on this film can be as detrimental to satisfactory seal operation as the physical characteristics of the seal itself. Shaft finishes which are too smooth cannot conduct or support an oil film when the seal lip pressures are too high. Conversely, a shaft which is too rough will produce mechanical damage of the seal lip by abrasion and heat due to rubbing without lubrication. That is, the roughness peaks become points of excess loading, resulting in breakdown of the relatively thin lubricant much the same as in journal bearing operation. Carbonization of both the seal lip and the lubricant causes shaft wear, thereby providing, at times, a compatible surface. However, during this *wear-in period* damage to the seal lip may be catastrophic.

Another effect of shaft surface finish which naturally follows is increased frictional torque with increased roughness, as shown in Figs. 1 and 2. These curves represent data collected during testing of two different rubber compounds. Also included in the

* Paper presented at SAE Annual Meeting, Detroit, Jan. 13, 1960.

quality control specifications

Reduce "LEAKERS"

data is the effect of seal lip pressure, which will be considered later. However, as lip pressures increase it is evident that oil films become thinner, resulting in higher torque values as shown. This information was obtained from the torquemeter, with the test shafts turning at 58 rpm and the seal lips fully lubricated. Of course, it must be understood that higher torque values would be obtained at higher speeds due to the increased rate of lubricant shear and abrasion of the seal lip.

Coupled with shaft surface finish may be the effects of shaft geometry on seal operation. Fig. 3 shows a typical trace of a production shaft. This particular trace was made on a Taylor Hobson "Talyrond" machine. A shaft of this geometry immediately points to another problem; that is, the ability of the seal to follow such a surface. Various degrees of shaft out-of-round can be tolerated depending upon the operating conditions, such as shaft speed, temperature, and viscosity of the sealant. At high shaft speeds, the recovery properties of the seal lip may be slow enough to allow a standing gap to form. A stiffening of the lip material due to extreme cold or heat would cause the seal to fail on a shaft such as this, but to function perfectly on other more nearly round shafts. The camming action of the high points require that seals of relatively higher lip pressures be applied. The resulting high loads due to inertia effects on the seal lip cause lubricant breakdown and associated high wear. Shaft run-out has the same effect as out-of-round. The degree of run-out which can be tolerated is again a function of shaft speed, temperature, and viscosity of the sealant.

Most steels used for shafting have microscopic voids which extend intermittently along the shaft, throughout its entire diameter. These defects occur in varying degrees in both hot and cold drawn steels, depending upon the composition and manufacturing variables. These flaws are readily detected by magna-fluxing and may be of the order of several microinches wide and several thousandths of an inch deep. The axial flaw shown in Fig. 4 caused the failure in a transmission seal application. It is believed that roll-bearingizing or roll-peening will close voids such as these. Some closure does occur as is shown in Fig. 5. However, flaking has occurred around the edges of the original void, which does

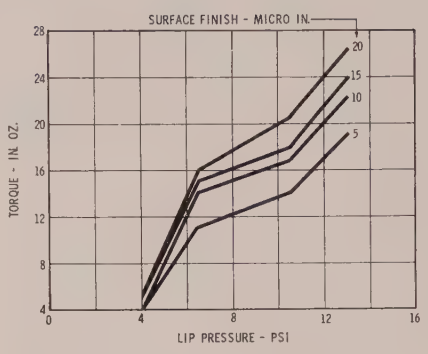


Fig. 1 — Torque versus lip pressure

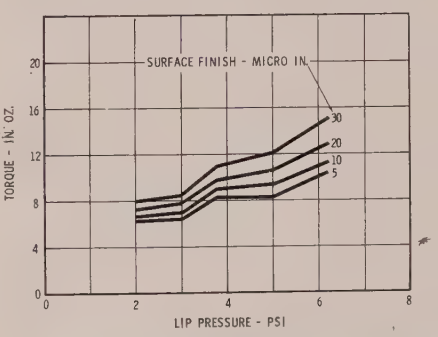


Fig. 2 — Torque versus lip pressure

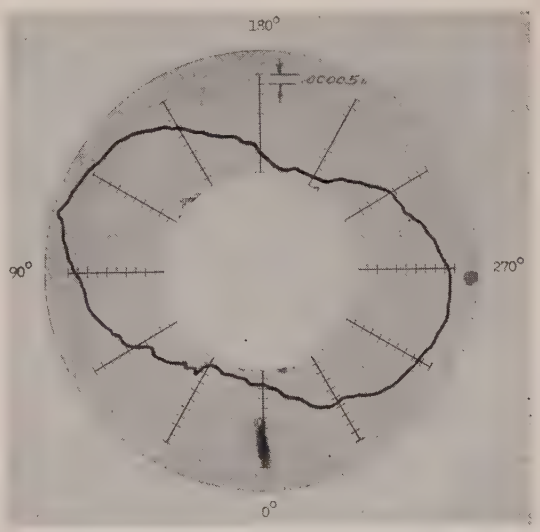


Fig. 3 — "Talyrond" trace of production shaft

not present a good running surface to the seal.

The effect of machining lead in a shaft surface can cause seal leakage by mechanical transfer of the liquid under the seal lip. Lead is defined as the helical machining grooves formed on the shaft surface as the machining tool moves along the shaft. The angle is controlled by the rate and direction of tool travel, and the rotational speed of the shaft. Fig. 6 shows a production shaft surface specified as 10 microin. and no lead. The seal which ran on this shaft leaked excessively. As a contrast, the shaft sample shown in Fig. 7 was prepared to the required specifications of 10 microin. and no visible lead. Since lead is an important variable, it is imperative that this be recognized in engineering testing, so that test shafts are used with the same lead and surface finish as the production shaft being investigated. There is a definite correlation between lead and surface finish, since considerable lead can be tolerated, providing the surface finish is fine enough to retain the sealant by a wiping action of the seal lip. This fact has been established by numerous tests to determine that seals will or will not leak depending upon the direction of shaft rotation. Apparently, the mechanism of leakage in this instance is the fact that the helical grooves are long enough to span the contact path of the seal lip. Thus, once the shaft is set in motion the lip material is unable to extrude into the groove as it did when stationary. If then the shaft is rotated in one direction, the sealant will be scooped into the groove and carried under the seal lip. However, when the shaft is reversed the helical grooves impart an axial component to the sealant away from the lip, and the seal will not leak. Here again, shafts with uncontrolled leads and surface finishes require the application of high lip pressure seals to machine the shaft during *run-in*.

Assembly

Various uncontrolled assembly conditions which result in nicks, scratches, burrs, and the like, both to the shaft and seal, are primary reasons for application failures. These are material handling problems and will not be discussed in this report. However, this factor is perhaps second in importance and must be controlled before corrective measures can be applied to quality control of seals.

Seal

Development and research work in lip seals during the past two years has shown that production variables in seal manufacture, and the lack of adequate inspection equipment for control of these variables, is the third primary cause for "leakers" which occur immediately or shortly after a product is put into service. In most instances, an adequate design can be made and satisfactory operation can be attained for a majority of seals in the same application. Those that do leak are perhaps less than 1% of the total, which in most cases can be directly attributed to physical characteristics resulting from production variables such as lip diameter, lip pressure, and eccentricity.

Seal Diameter Control

The seal lip diameter, as determined by the amount of trim interference and the degree of compression applied by the garter or leaf spring, if accurately measured can be correlated to seal leakage. A rash statement to say the least; however, with some qualification a large percentage of seal leaks can be directly attributed to diameter variations. A diametral analysis of seals in several critical applications indicates a wide distribution as a result of trim variation and spring quality. The effects of trim interference on the ability of a seal lip to perform properly are shown in Figs. 8 and 9. Fig. 8 shows a series of tests performed on a number of seals manufactured from a laboratory batch of Buna compound. However, instead of merely trimming the seals to the various interference fits, the single cavity molds were constructed to give the desired interferences with the trim line held constant in relation to the spring centerline. This was done to to avoid placing the spring pressure at various positions in the head section as would be necessary with a single mold. You will note that seals having zero and 0.030-in. trim interference display an extremely low frictional characteristic, while those at 0.060- and 0.090-in. trim interferences cause the material to affect an increasing frictional drag on the rotating shaft even though the lip pressures were held constant by varying the garter spring length.

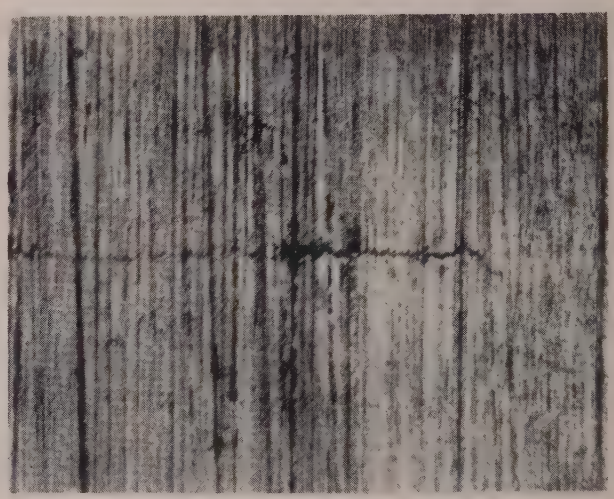


Fig. 4 — Axial flaw at shaft surface

Fig. 5 — Axial flaw at shaft surface, after burnishing

This wide variation in torque is explained by the fact that rubber-base materials display a curious physical property known as "Joule effect." This is a thermoelastic effect which causes rubber to shrink when it is subjected to tensile stress and heat simultaneously. Conversely, when rubber is placed under compression and exposed to elevated temperatures it will grow. Of course, this phenomenon occurs in varying degrees, depending upon the compounding and various other physical properties of the rubber. A complete knowledge of this property is not available at the present time but it is evident that it exists in production seals, as shown in Fig. 9. These seals were selected with identical physical characteristics except for trim interference. They were run on test in a calorimeter type of test head which recorded heat input to the lubricant as a function of operating speed. Here the effects of the greater trim interference are clearly evident. Suffice it to say that diameter control is important from the standpoint of trim, as shown in each of the seal applications investigated to date.

Referring again to Fig. 8, each of the seals used in this series of testing had identical lip pressures and the only variable was the trim interference which varied the amount of shaft interference from 0.000 to 0.090 in. Not only was the variation in torque investigated, but also lip flutter which is perhaps better described as stick-slip. This was detected by means of magnetic pickups placed in close proximity to the garter spring as shown in Fig. 10. Due to wrapping of the seal material around the shaft and a sudden release, a radial component is imparted to the garter spring which varies the flux density of the magnetic pickup, thus, by proper amplification and integration of the signal, a displacement trace of lip flutter could be produced on an oscilloscope. Contrary to the common belief that during stick-slip leakage occurs, the opposite effect is true. During stick-slip, high torque occurs which results in a sharp rise in temperature at the seal rubbing face. Heat decomposition and mechanical abrasion of the seal material will eventually subside and leave the lip severely damaged and unable to effect sealing. It is also interesting to point out that stick-slip can occur in one segment of a seal, resulting in damage to the lip at this point only. Such wear is caused by variations in lip pressure and seal section. Therefore, seal leakage can occur as a direct result of diameter variations.

Compatible diameters for a seal in a given application will be determined by the seal design, type of rubber compound, and shaft surface finish, so that each must be evaluated separately. In other words,

Fig. 6 — Production shaft — 10 microin.?

Fig. 7 — Ten-microin. surface, no lead

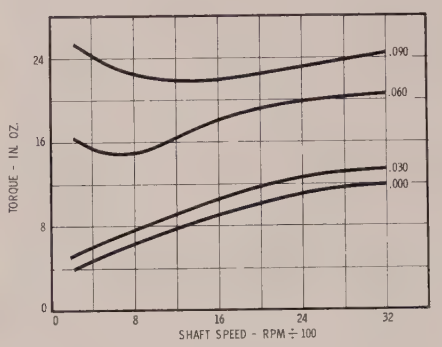


Fig. 8 — Effect of trim interference on torque

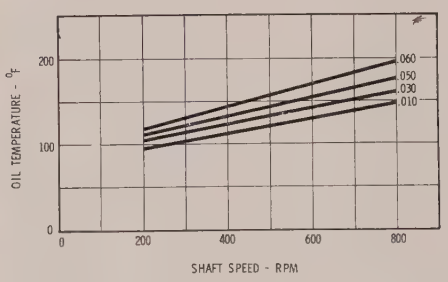


Fig. 9 — Effect of trim interference on seal temperature

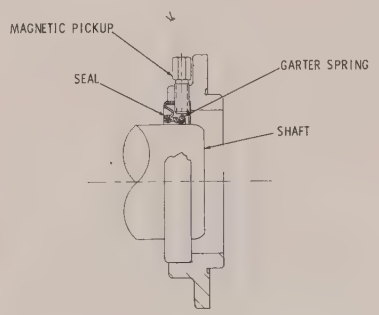


Fig. 10 — Magnetic pickup installation

diameters which are compatible for one seal cannot be applied to another seal design, manufactured for the same application. The mention of surface finish having an effect on compatible seal diameters has been substantiated by control testing and application of these controls to production sealing problems. This is perhaps a point of controversy, but the fact remains that a smooth shaft, 5 microin. will require much closer selection of seal diameters than the same shaft finished to 15 microin. This is perhaps a direct result of the smooth shaft's inability to conduct and retain microscopic pockets of lubricant under the seal lip, resulting in high torque for tight seals. This has been demonstrated in the laboratory by running a seal with high trim interference and moderate lip pressure against a shaft with 2-microin. surface finish. By means of torque measurements and the magnetic pickup described earlier, the seal was observed to be in a state of stick-slip and high torque. This condition could be immediately stopped by moving a shaft section having a 10-microin. surface in contact with the seal lip. Furthermore, field experience shows that two identical production seal applications having different surface finishes require a seal diameter range of 0.030 in. for the 5–8 microin. shaft and 0.050 in. for the 15–18 microin. shaft.

Examination of frequency size distribution patterns of a number of critical seal applications shows a wide variation for identical seals, as illustrated in Figs. 11 and 12. Fig. 12 shows diameter variations below 1.815 in. and above 1.845 in. as designated by the shaded areas in the batches sampled at three-day intervals from production. Initially, over 7000 hr of closely controlled test work on these seals showed that satisfactory diameters for this application range between 1.815 and 1.845 rather than that included in the frequency distribution curve. A more detailed chart of this distribution curve (Fig. 13) shows the high precentage of "leakers" which occur outside the diameter limits, while those within the range showed leakage rates well below the 0.100 gram/50 hr, set as a standard. Repeat of this testing on a similar batch of seals substantiated the initial findings which indicated a necessity for establishing diameter limits at the designated points. This type of diameter control has been applied in seven additional installations with equally gratifying results.

In order to control seal diameters in high volume production, it was necessary to develop an instrument now known as the "Sealector." This is an air

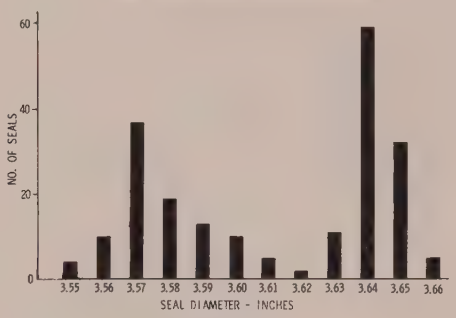


Fig. 11 — Frequency size distribution of production seal design

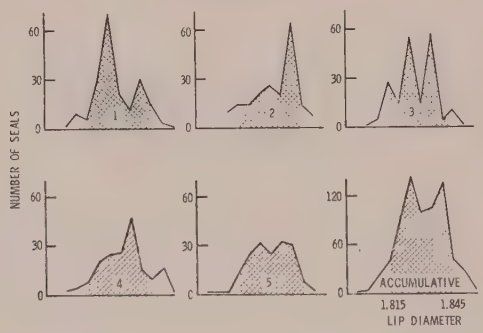


Fig. 12 — Frequency size distribution of production seal design

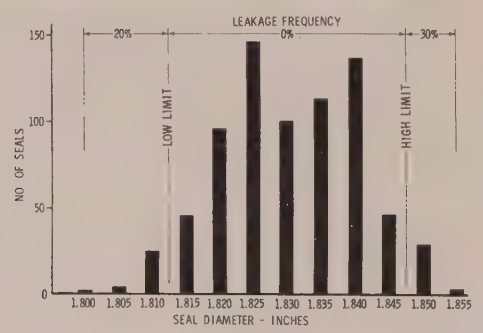


Fig. 13 — Diameter limits for transmission application

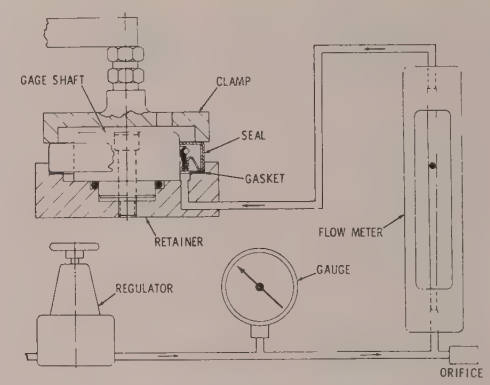


Fig. 14 — "Sealector" schematic

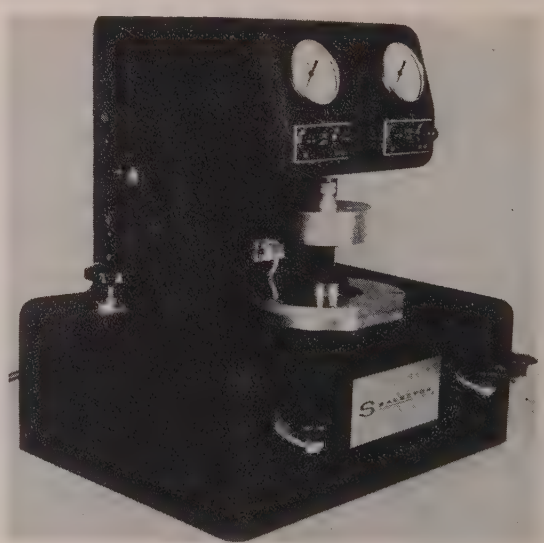


Fig. 15 — "Sealector" — semi-automatic

gage, designed to indicate seal lip diameter by comparing (at standard pressure) mass flow. A schematic diagram of this system is shown in Fig. 14. If the seal lip is concentric to the shaft and has a diametral clearance of several thousandths of an inch, then a flow of air will occur through this orifice which will be on the order of 10,000 cc min. The accuracy of this system is such that seal lips can be gaged in increments of 0.005 in. on the diameter, with the correlation of results between various operators being high. Fig. 15 shows a semiautomatic "Sealector" which uses a system of electronics and colored lights to indicate seal acceptability on two diameter limits. This instrument has been operated at rates up to 25 pieces/min so that it is possible to check minimum and maximum diameters of all seals in any given application.

Seal Lip Pressure

Seal lip pressure is a term used to designate the radial force exerted by the lip on the shaft as a result of trim interference, rubber hardness, section uniformity, eccentricity, and spring quality. Here we have a number of variables, any one of which could effectively cause a seal to malfunction. With manufacturing controls presently applied to seal manufacturer, these can be held within relatively close limits. However, let us examine the effects of lip pressure on seal operation. Fig. 16 shows a series of data obtained from tests performed on seals manufactured from a laboratory batch of Buna material and trimmed to 0.030-in. shaft interference. These were fitted with garter springs to produce lip pressures as designated. Subsequent testing under the same operating conditions showed that those seals having the higher lip pressure produced excessive torque and heat.

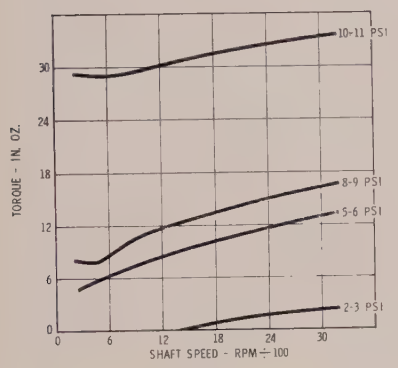


Fig. 16 — Effect of lip pressure on running torque

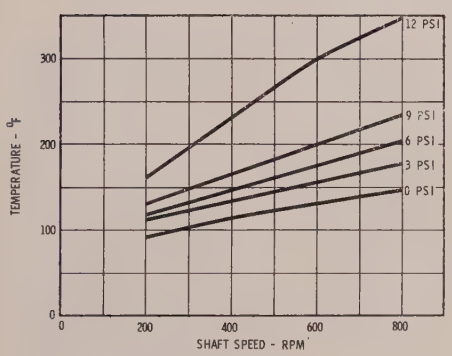


Fig. 17 — Effect of lip pressure on seal temperature

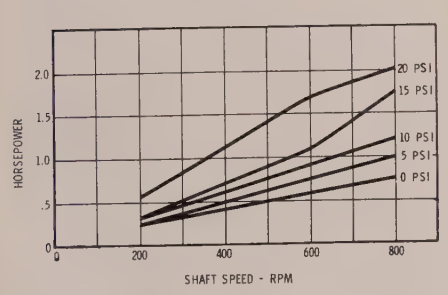


Fig. 18 — Seal lip power requirements versus lip pressures

Previously, we mentioned a study in which magnetic pickups were used to detect seal lip movement associated with stick-slip as a result of seal diameter. Actually, seals which fluttered due to diameter variations were initiated into this unstable condition by a change in lip pressure during operation, as a result of "Joule effect." Apparently, the mechanism of stick-slip is a direct breakdown in the lubricant film between the seal lip and shaft as a result of high temperatures. The reduction in lubricant viscosity as a result of these temperatures allows the seal lip to contact the shaft and be excited into stick-slip because of increased torque. It is interesting to note that stick-slip occurs at a relatively high frequency, for instance, 1280 cps observed in one series of testing. Also this characteristic is independent of shaft speed; in other words, the seal lip will flutter at this frequency throughout an entire range of speeds from several hundred to several thousand shaft rpm. The fact remains that flutter is initiated by change in lubricant temperature within very critical limits.

Recent studies have revealed that seal lip temperatures in excess of 200 F differential can exist above the surrounding lubricant, as demonstrated in Fig. 17.

These tests were conducted on a 7-in. diameter seal subjected to air pressures as indicated. Although completely submerged in oil, which never exceeds 160 F, the seal lip temperatures exceeded the limits of the material. These temperatures were obtained by imbedding thermocouples in the shaft surface and grinding them flush. By holding the shaft stationary and rotating the seal lip over the thermocouple, the running temperature at the contact face was recorded under the various operating conditions. The small change in lip pressure due to centrifugal force on the rotating seal lip was compensated for by applying air pressure. Lubricant just sufficient to cover the seal lip and shaft was supplied at a constant temperature to simulate actual conditions. By means of a strain beam attached to the stationary shaft, which was supported in ball bearings, continuous horsepower measurements were made along with temperature (Fig. 18). This graph readily demonstrates the effects of lip pressure on the operating characteristics of this particular seal. In numerous applications where the sealed shaft is not properly manufactured in regard to surface finish, seals have been applied with extreme lip pressures to do the job of machining. This practice, to say the least, is highly unsatisfactory since control of each of the aforementioned factors can be such that the seal will perform satisfactorily without shaft contact. In other words, with proper control of lip pressure, lip diameter, and shaft surface finish, a seal can be made to operate as a precision-fitted bearing and no apparent con-

tact with the shaft. Tests of this type have been conducted for periods in excess of 100 hr with maximum sealing and no shaft contact. These, of course, are laboratory controlled conditions but it does demonstrate that extended life and perfect sealing can be obtained by control of lip pressure variables. Test data show that although seals are manufactured and applied with high lip pressures, after less than 100 hr of operation they lose all pressure in excess of that required for satisfactory operation, either by compression set, wear, or both. The seal having gone through this physical change may or may not operate without leaking. In any event, the seal and the shaft are being machined to fits more suitable for the operating conditions imposed.

Earlier it was stated that rubber hardness contributes to lip pressure effects. In addition, it contributes to the stability of the seal section so that control of this factor in seal production is important. For instance, a range of rubber hardness from 69 to 84 Shore 'A' affects the operating temperature of the seal rubbing face tremendously, as shown in Fig. 19. This chart shows lip temperature variations for two 7-in. diameter seals operating with their manufactured lip pressure, and an added external air pressure of 20 psi. In control testing to investigate this effect, it was discovered that seals with the lowest running temperatures corresponded to those having the lowest rubber hardness. Examination of the wear patterns showed a rubbing path width of 0.180 in. for the 69 Shore 'A' material, while that in the range of 84 'A' was on the order of 0.030–0.040 in. Going back to the bearing theory, let us examine the wear paths in respect to the lubricant film. As shown in Fig. 20 those seals with low hardness deformed sufficiently under load to contact the shaft in a wide path sufficient to build up a hydrodynamic film of lubricant. This film creates an opposing pressure head to counteract the lip pressures and reduce material contact, consequently the lower temperatures. Seals having high hardness and narrow wear paths cannot support this type of film under high loads, so that the existing film is broken to cause material rubbing and high temperatures. However, it must be pointed out that seals operating with a wide contact and the hydrodynamic film will invariably leak due to the greater running clearances formed.

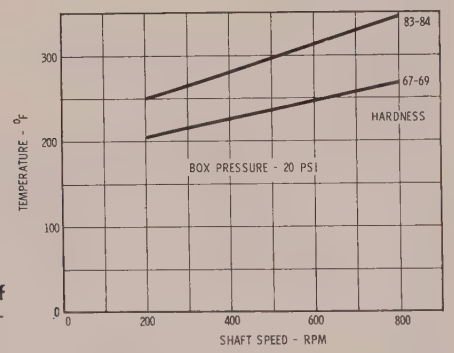


Fig. 19 — Effect of lip hardness on running temperature

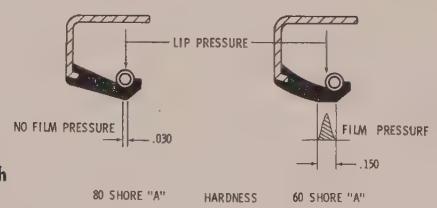


Fig. 20 — Wear path of lubricant film

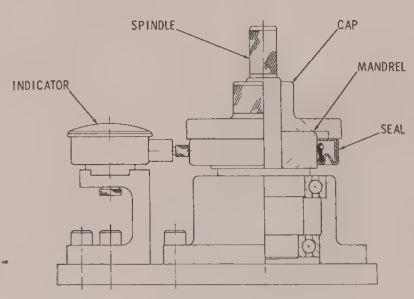


Fig. 21 — Eccentricity gage

A modification of the "Sealector" can be used to control manufactured lip pressure variables. This gage uses the same circuitry as the diameter checking gage, except for the fact that a mandrel of shaft size is used along with a cam-operated regulator. The seal is clamped over the shaft and the air pressure raised by means of the cam until the seal lip expands away from the shaft sufficiently to cause a small flow at which time the lip opening pressure is read. The pressure at which this flow occurs approximately indicates the radial pressure that the combination of above mentioned factors causes the seal lip to apply to the shaft during operation. This observation holds true for concentric seals with uniform molded sections and rubber compound. In a production seal, this lip pressure reading indicates a combination of the characteristics of the seal section, eccentricity, rubber hardness, and spring assembly. In other words, radial pressure at any point on a seal lip is not uniform and a pressure check will show the pressure as being exerted by the weakest part of the lip. A normal production distribution of lip pressure for a given seal design will vary from one-half to several pounds and still give satisfactory operation if diameter limits are maintained.

Seal Eccentricity

The eccentricity of the seal lip in respect to its case is extremely important with regard to the effect on lip pressure. For instance, a 2-in. diameter seal will show a change in lip pressure of 1 psi when placed eccentric 0.005 in. Several causes for seal lip eccentricity can be element displacement in a mechanically builtup assembly, section change due to poor mold register, and the trim operation. The lack of eccentricity information in engineering and production acceptance testing has resulted in a wide scatter of data which could only be attributed to the rubber compound variables. Seal eccentricity is perhaps responsible for the major portion of failures which occur in testing, since the effect is not recognized when assembling the seal in a test fixture which has built in eccentricity. For instance, most all seal test equipment is constructed to provide at least 0.010-in. eccentricty as one of the test criteria. It is, therefore, obvious that a seal which has 0.010-in. eccentricity could be assembled and tested with a total of 0.020 in. resulting in a seal failure, or it could inadvertently be assembled with zero eccentricity and pass the test as a good seal.

Eccentricity of the seal lip can also be affected by

the application of the garter spring. Applying a spring by hand will pull the seal lip eccentric and if the tension is sufficient to cause the coils to bed into the rubber, the seal will operate with this assembled eccentricity for extended periods and eventually fail. This is especially severe in large diameter seals so that it is necessary to assemble the springs by rolling them into the spring groove over a mandrel. However, even this type of assembly does not assure proper application since the spring may not leave the mandrel uniformly.

In very critical applications concentricity values are measured by means of a gage such as shown in Fig. 21, where the seal is placed over a teflon-coated shaft by means of the guided cap. This cap is then removed and the seal allowed to seek its manufactured center, which is measured by means of the indicator. For less critical applications, this parameter can be controlled by lip pressure and diameter measurements on the "Sealector."

DISCUSSION

"Joule" Effect Does Not Exist In Shaft Seal Design

— Warren Rasmussen
Chicago Rawhide Mfg. Co.

INITIALLY, I would like to discuss the experience our company has had relating to the use of the Sealector device which, as Mr. Dega indicated, is a gage to control seal diameters in high volume production. We find that within certain limitations the Sealector affords both the seal manufacturer and the seal user an excellent means of judging consistency of product. One thing which I believe should be understood regarding the Sealector is that it measures not the rubber ID of a seal alone, but rather a combined function of the rubber ID and the eccentricity within the seal itself. As the authors carefully stipulated, a seal which has a lip concentric to the shaft and a dimetral clearance of several thousandths of an inch when placed on the Sealector will result in an airflow through the Sealector orifice in the magnitude of 10,000 cc/min. Now, if the seal lip is not concentric to the shaft, but rather sufficiently eccentric, you will no longer have a uniform orifice opening. This will result in an increased flow of air at the constant test pressure. The net result of such increased flow is to cause acceptance of seals which should normally be rejected because their ID's are too small and conversely cause rejection of seals with ID's close to, but not above the maximum limit. We believe most people familiar with seals, their design, and their usage will recognize quickly that internal eccentricity in a seal is undesirable. Tossing out seals that are near the high limit because of eccentricity is probably a plus factor for the Sealector. However, acceptance of seals with very small lip diameters because of large amounts of eccentricity in the seal certainly is not desirable.

With regard to Sealector inspection techniques, we have seen an expanded use of this equipment in the past six to eight months. Unfortunately, not all users recognize that each seal must be considered as an entity within the framework of its operating environment to establish proper control limits on diameter and pressure. As the authors indicate, this is true because compatible lip diameters and pressures are a function of a complex of variables including specific seal design, compound, shaft finish and eccentricity, medium being sealed, and other operating conditions. It would certainly be incorrect procedure to take limits established on one seal and apply them directly to another. In many cases it would also be dangerous to take limits established for one application and apply them to the same seal in another application.

There is, finally, one point made by the authors with which we take exception. This involves their explanation of the wide variation in torque relative to variations in seal lip trimmed interference, as depicted in Figs. 8 and 9. As you will recall, all of the seals were equated in terms of lip pressure and differed functionly only in the amount of molded-in interference which varied from 0.000 to 0.090 in. The authors have attributed the variations in frictional torque and heat generation to the "Joule" effect which they denote as a thermal elastic effect which causes rubber to contract when subjected to a combination of tensile stress and heat simultaneously. Some years ago, my company, and I believe a majority of companies in the seal industry, subscribed to the existence of the "Joule" effect in seals, and in many instances went to great extremes to design seals without molded interference so as to circumvent the "Joule" effect. Based on a series of tests which we have conducted on an Instron Tensile Tester, we confidently state that there is no "Joule" effect that need be considered in shaft seal design. By this comment I am not attempting to infer that the "Joule" effect as such does not exist. It certainly does, and we have demonstrated such in many natural rubber products. However, we have conducted extensive evaluations of a variety of synthetic rubber materials representative of the vast number of compounds used in shaft-type seals and have yet to find one of these synthetic rubber compounds which exhibits a phenomenon remotely similar to the "Joule" effect.

When a section of a natural rubber rubber-band is deflected in tension and then subjected to increases in temperature, the load increases if the deflection is held constant. Conversely, every test we have ever conducted along this same line on synthetic rubber materials ranging from Buna-N's to Viton have indicated just the opposite effect; namely, with increasing temperature, a decrease in load.

In as much as I am not aware of the details of the investigations which developed the data underlying Figs. 8 and 9, I unfortunately cannot offer a satisfactory explanation as to the cause of the particular effect shown. Nonetheless, I am certain that, this reaction cannot be attributed to the "Joule" effect and additional investigations appear necessary to determine the true cause.

ORAL DISCUSSION

Reported by E. A. Slindee
International Harvester Co.

J. Born, Chicago Rawhide Mfg. Co.: How do you measure oil film thickness?

Mr. Symons: It was not measured. We would like to know how. It was approximated by varying the shaft finish in 5 microin. steps and observing when the oil film broke down.

Wayne Hartman, J. I. Case Co.: Is the Sealector available for purchase?

Mr. Symons: Yes, contact Tom Vandergift from the Special Problems department of G. M. Research Laboratories.

Mr. Born: Does the lip shape have any affect on the seal?

Mr. Symons: Yes, both shape and lip pressure are important.

R. Cellitti, International Harvester Co.: Please elaborate on the box pressure. Why do you increase it?

Mr. Symons: This is a standard railroad seal test and is the actual condition in a railroad box.

T. Anderson, Chicago Rawhide Mfg. Co.: Do different contact widths at the lip with various equivalent lip pressures leak the same amount?

Mr. Symons: We have no data on this. However, over about 0.100 in. width, the hydrodynamic effect causes leaks. Therefore, the contact width does have to be controlled within certain limits.

•Oxidation Stability

•Shear Stability

•Rubber Swell Properties

H. E. Deen and C. M. Stendahl
Enjay Laboratories

MUCH has been written about the problems of formulating automatic transmission fluids for passenger-car use. These fluids have been aptly described as the most detailed and precise products made by a petroleum refiner. It is the purpose of this paper to discuss three important aspects involved in preparing automatic transmission fluids: their oxidation stability, shear stability, and rubber swell properties. New techniques developed by our company in studying these properties are described. Data have been obtained on a variety of fluids using these techniques and the results related to existing and possible future specification tests.

Oxidation Stability

Sludge and varnish are two well-recognized end products of the oxidation of hydrocarbon lubricants. They are highly undesirable in all lubricant applications. In particular, they are unwanted in precision-built automatic transmissions where they can deposit in critical areas and cause malfunction or damage.

Since all present-day ATF's oxidize under normal usage, they will eventually reach a point where sludge and varnish begin to form. To prevent this, transmission manufacturers recommend that fluids be changed at periodic intervals. The recommendations vary, depending on the individual manufacturer's experience on normal transmission operating temperatures and ATF oxidation stability requirements. The trend in automobile and automatic transmission design has resulted in higher temperatures, which means more severe operating conditions. Consequently, the oxidation stability properties of ATF's must be improved.

The importance of oxidation stability to automatic transmission manufacturers is emphasized by the various equipment tests which have been set up to measure this property. Four of the best-known tests are described in Table 1.

These tests are recognized to be quite severe, and fluids with good oxidation stability are required to pass them. Future transmissions, however, are expected to require even greater oxidation stability than the tests can guarantee. In anticipation of this, the automotive companies have introduced

A NEW TEST is described for studying the oxidation stability of automatic transmission fluids (ATF). The test shows an excellent correlation with transmission oxidation tests and points out the importance of time as a variable in such studies.

Carefully controlled automobile dynamometer tests have been used to study the shear stability of ATF's. Data are presented showing a comparison of driving conditions, transmissions, and V.I. improvers on shear stability. Results are related to the 50-hr Hydra-Matic durability test.

The poor reproducibility of rubber swell measurements on commercial transmission seals is due largely to differences in the rubber compounds. A great improvement in the reproducibility may be made by taking into account the specific gravity of the rubber sample.*

# of Automatic Transmission Fluids

two new tests which are described in Table 2.

It is quite obvious that running one of these transmission tests to measure oxidation stability is both expensive and time-consuming. It is very desirable, therefore, to have a bench test which can predict the performance of fluids before running them in the more expensive equipment tests. A number of bench tests have been used with varying degrees of success by the industry. Foehr and Calish[1] described their results with a slight modification of the Allison Type C bench oxidation test. Kieninger and Retzloff[2] used the Sohio Polyveriform lubricant tester. Quigg[3] used a modified Norma Hoffman test. Others that have been used are the Indiana oxidation test and a modified MacCoull-Ryder test. A description of these tests is shown in Table 3.

Although the conditions employed in the tests vary quite widely, one thing appears common to them all: Oxidation conditions are greatly exaggerated in order to shorten the test time. Such methods as the use of copper catalyst, very high test temperatures, air blowing, and pressurizing with pure oxygen are employed. Another thing common to all but the Indiana test is that the test time is fixed, and results are analyzed only at the end of the test.

We studied two of these tests and obtained very unsatisfactory results. Both the Type C bench oxidation test and the modified MacCoull-Ryder test showed no correlation with Powerglide L-39 transmission tests. A program was, therefore, undertaken to develop a new test.

Enjay Bench Oxidation Test

At the time the Enjay test was developed, the new transmission tests described in Table 2 had not been introduced. However, it was apparent that tests more severe than the Powerglide test at 275 F were being considered by the automotive industry. The objective of our laboratories was, therefore, to develop a screening test for an expected future transmission test. A Powerglide transmission test run at 300 F was selected as a likely prospect for the future.

The bench test was developed with the following concepts in mind:

1. Oxidation conditions should be as similar as

Table 1 — Current Transmission Tests for Measuring Oxidation Stability of Automatic Transmission Fluids

Transmission Type	Power-glide	Merco-matic	Hydra-Matic	Torque-flite
Year	1955	Current	Current	1957
Test Conditions				
Sump Temperature, F	275	275	275	275
Time, hr	300	300	300	300
Speed, rpm	1750–1850	2150 ± 50	2000	1750
Shift Cycles	No	No	Yes	No
Aeration	No	No	Yes	No

Table 2 — New Proposed Tests for Measuring Oxidation Stability of Automatic Transmission Fluids

Transmission Type	Powerglide	Mercomatic
Year	1955–1957	Current
Test Conditions		
Sump Temperature, F	275	300
Time,hr		
Maximum hr Time	240	300
Intermediate Inspections	24-hr intervals after 144 hr	100 & 200
Speed, rpm	1750	2150 ± 50
Shift Cycles	No	No
Aeration	Yes	No

* Paper presented at SAE National Fuels and Lubricants Meeting, Chicago, Oct. 29, 1959.

[1] "Correlation of Bench Oxidation Tests with Transmission Test Performance of Automatic Transmission Fluids," by E. G. Foehr and S. R. Calish. Paper presented at SAE Annual Meeting, Detroit, January, 1959.

[2] Discussion of Foehr-Calish paper, by J. Kieninger.

[3] "Screening Tests for Automatic Transmission Fluid, Type A," by H. T. Quigg. Discussion of Foehr-Calish paper, footnote 1.

possible to those present in the transmission.

2. Time should be included as a test variable.
3. The test should be simple and inexpensive.

In the Powerglide L-39 transmission test, the amount of air contacting the fluid is limited. It was, therefore, decided to conduct the proposed bench tests in a closed oven to simulate this limited air-fluid contact. To simulate the constant agitation of the fluid in the transmission, a stirring mechanism was built into the oven. The equipment was designed to evaluate 11 fluids simultaneously. The fluids were placed in glass beakers, and agitated with glass stirrers. For a catalyst, a section of a bronze thrust washer from a Powerglide transmission was selected. A constant temperature of 300 F was used, and the maximum test time was 300 hr. Thus, from the design standpoint, the equipment satisfied the first concept.

To implement the second concept of time being a variable, another modification was made in the oven to permit sampling of the fluids during the test without opening the oven door. This consisted of drilling holes in the top of the oven and inserting tubes which were capped inside the oven with spring oil caps such as those commonly used on electric motors. The tubes were positioned directly above the glass beakers so that long syringe needles could be inserted through the top of the oven and oil samples withdrawn. Fig. 1 is a schematic diagram of the test. Fig. 2 shows the exterior view of the ovens.

The initial work with the test included a study of a number of variables, such as viscosity increase of the fluid with time, sludge and varnish formation on the stirrers and in the beakers, and analysis of the used oil for sludge and acid number. It was soon found, however, that the following simple procedure gave the best indication of the oxidation stability of the fluid. The procedure consists of sampling the fluid at 24-hr intervals, filtering 15 drops of it through a small filter paper, and visually rating the paper for sludge deposition. The results are quickly available and a conclusion as to the condition of the fluid is readily apparent to even untrained personnel. This met the requirements of the third concept — simplicity, and is the only technique that is now used to follow the oxidation stability of the fluid. Table 4 summarizes the test procedure. A detailed description of the procedure is given in Appendix I.

The results of the test as it is now run are best illustrated in Fig. 3. This is a photograph of the filter papers from tests on a number of commercial automatic transmission fluids. The most striking

Table 3 — Bench Tests for Measuring Oxidation Stability of Automatic Transmission Fluids

Type of Test	Allison Type C	Sohio Polyveriform	Modified Norma Hoffman	Indiana Oxidation	Modified MacRyder
Test Conditions					
Time, hr	150	116	300	Indefinite	48
Inspections During Test	No	No	No	Yes	No
Temperature, F	300	300	210	341	275–325
Air Injection	Yes	Yes	Pressurized with O_2	Yes	No
Mechanical Stirring	No	No	No	No	Yes
Catalyst	Iron & copper wire	Steel & copper	Iron & copper wire	No	Copper
Criteria for Rating Fluid					
Used Oil Conditions					
Acid no. increase	No	Yes	No	No	Yes
Insolubles	No	No	No	Yes	Yes
Viscosity change	Yes	Yes	No	Yes	Yes
Equipment					
Sludge	No	Yes	No	No	Yes
Varnish	No	Yes	No	No	Yes
Other	No	No	Pressure drop of O_2	No	Visual inspection & weight loss of copper catalyst

Table 4 — Enjay Bench Oxidation Test Conditions

Test Conditions	
Time, hr	300, maximum
Inspection during test	Yes, filter paper analysis of fluid every 24 hr
Temperature, F	300
Air Injection	No
Mechanical Stirring	Yes
Catalyst	Portion of bronze thrust washer
Criterion for Rating Fluid	Time to first formation of sludge in filter paper test

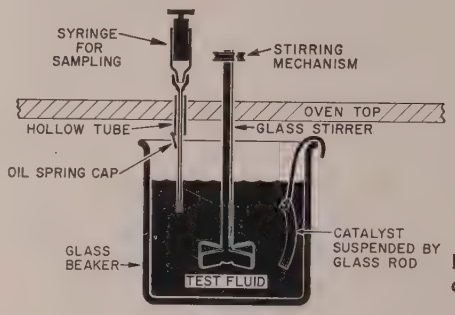


Fig. 1 — Schematic diagram of Enjay Bench Oxidation test

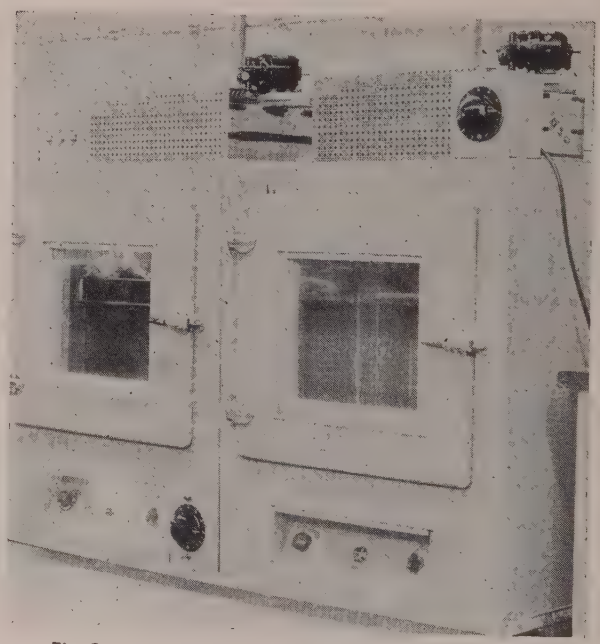


Fig. 2 — Exterior view of ovens used for oxidation test

thing to be noted about these papers is the sharp break in the oil with time. The terminology "break point" is used to describe the condition where sludge particles first appear on the filter paper. It may be seen that the first oil (A) had broken before the first sample inspection at 75 hr. Oil B showed a sharp break at 99 hr, while Oil C, which is a CRC reference fluid for the Powerglide L-39 test, showed a sharp break at 123 hr. The best fluid in these tests, Fluid H, showed a break point at 219 hr.

For correlation studies, this break point in the bench test was related to a general description of the performance of transmission fluids in the Powerglide L-39 test. The correlation is shown in Fig. 4. It may be seen that the oils which were poor in the L-39 test broke quite early in the Enjay bench oxidation test, usually in less than 100 hr. Fluids rated satisfactory or good in the transmission test broke at close to 120 hr, while fluids which were excellent in the transmission test did not break until close to 200 hr in the bench test.

More recent correlation studies were carried out between the bench test and the Powerglide test with air injection. A good reference fluid supplied by General Motors Research was run in both the bench test and the Powerglide test with air injection. The oxidation stability of the fluid in the transmission was followed using the same filter paper technique employed in the bench test. Fig. 5 shows the filter papers from the two tests. The papers have been rated on a standard demerit scale where 0 indicates no sludge particles and 10 indicates so much sludge that the sample is unfilterable. It may be seen that the sharp break point that occurred in the bench test also took place in the transmission. In the bench test, the break point was 192 hr. In the

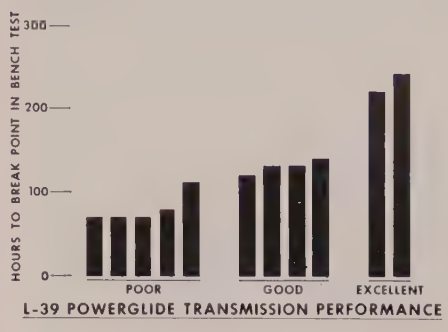


Fig. 4 — Correlation with CRC L-39 Powerglide transmission test

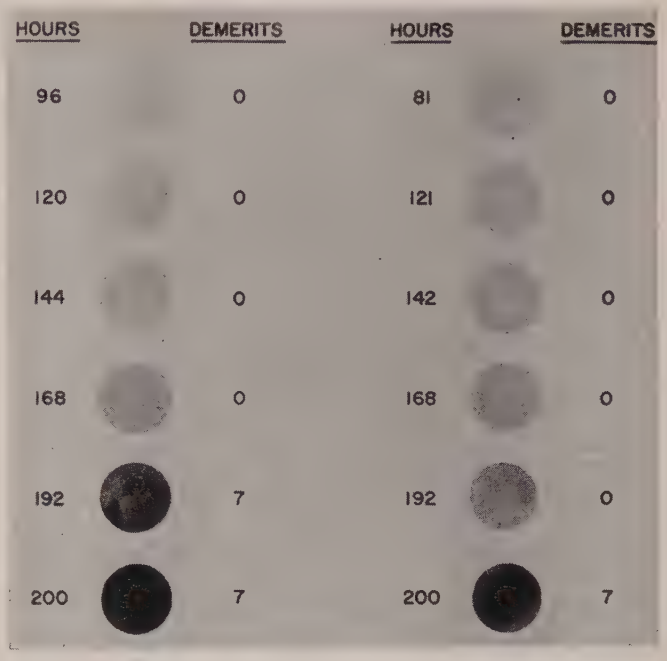


Fig. 5 — Comparison of Bench Test (left) with Powerglide Test with air injection (right) using good reference fluid

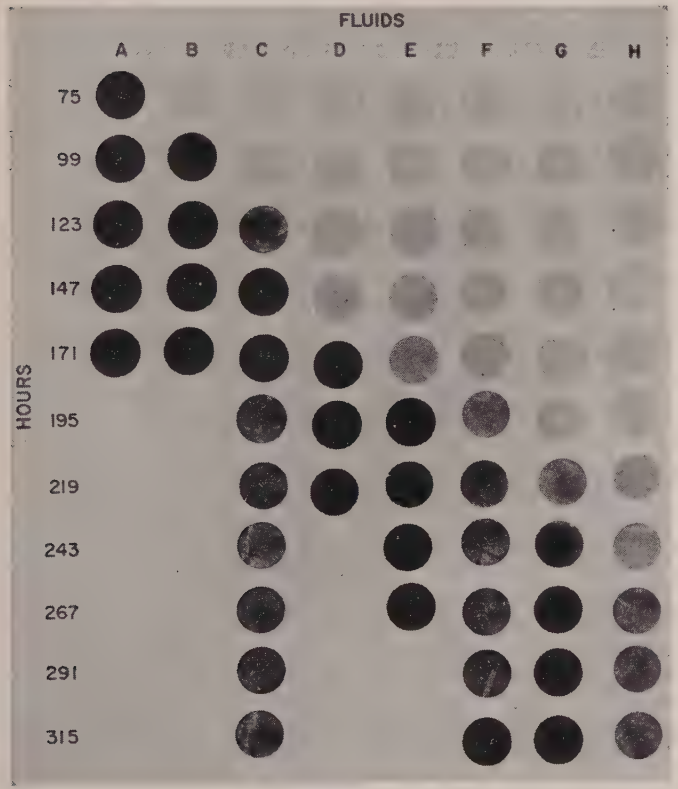


Fig. 3 — Filter paper ratings of commercial automatic transmission fluids in Enjay test

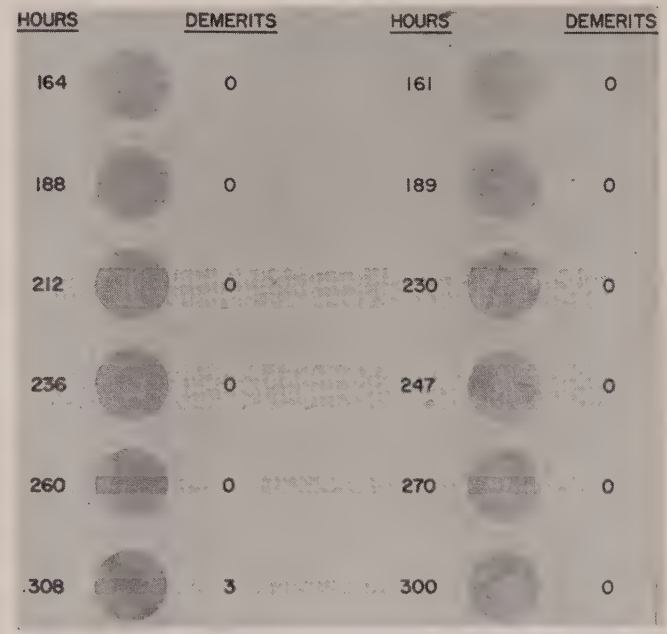


Fig. 6 — Comparison of Bench Test (left) with Powerglide Test with air injection (right) using excellent quality fluid

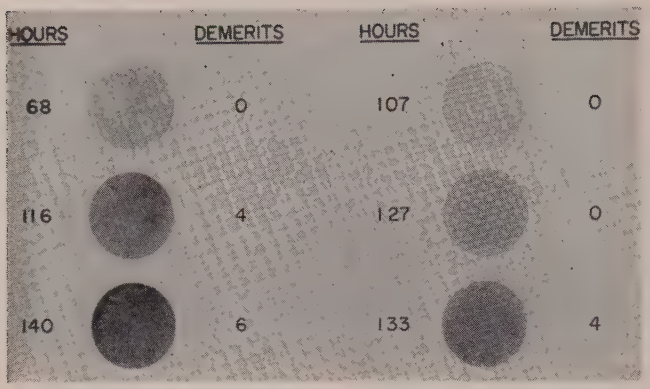


Fig. 7 — Comparison of Bench Test (left) with Powerglide Test with air injection (right) using poor quality fluid

transmission, the fluid was good at 192 hr but broke at 200 hr, as indicated by the filter paper test of the used oil. Inspection of the low and drive valve body at this point showed trace sludge and light varnish formation. Thus, the break point of the fluid as specified in the procedure for the Powerglide test with air injection[4] was predicted by the bench test results.

A fluid which gave excellent results in the bench oxidation test also gave excellent results in a Powerglide test with air injection. This is illustrated in Fig. 6. It may be seen that the fluid did not break until 308 hr in the bench test and showed no break in the 300-hr transmission test. At the end of the test, the transmission was found to be free of sludge and varnish deposits. Fig. 7 shows a poor fluid which broke at 116 hr in the bench test and showed a sharp break at 133 hr in the Powerglide test. Once again, inspection of the low and drive valve body at 133 hr showed sludge and varnish formation, indicating the fluid had broken in terms of the test procedure.

Altogether eight data points are available comparing the Powerglide test with air injection and the Enjay bench oxidation test. The transmission tests have been conducted both at our laboratories and outside laboratories on fluids containing a variety of additives. The data are plotted in Fig. 8 and show that an excellent correlation exists. In all cases, the break point in the transmission test is taken as the first formation of sludge or varnish in the low and drive valve body. In the bench test, the break point is determined by the filter paper technique.

No sound judgment on the value of a test can be made without some idea of its variability. A limited amount of data has been obtained which indicates the repeatability of the break point in the Enjay bench test is quite good. Fluids are analyzed for the break point at only 24-hr intervals and a fluid is considered to have broken when the first appearance of sludge particles on the filter paper is noted. Thus, depending on the relationship of the time of sampling to the actual break point, results could vary by as much as 24 hr. The available data on the test repeatability are given in Table 5. It may be seen that the standard deviation is 13 hr for fluids in the excellent oxidation stability range of 200–300 hr. This accuracy has proved entirely satisfactory for the work that has been conducted to date. Obviously, if desired, the standard deviation could be reduced by employing more frequent sampling of the fluid than the present 24-hr frequency.

The Enjay bench oxidation test is being set up at other research laboratories; so at some future date, information on the reproducibility of the test may also be obtained.

The main conclusion that can be reached from these studies is that time is a very important variable in studying the oxidation stability of ATF's. Once the oxidation inhibition properties of a fluid have been spent, sludge and varnish formation proceed at a rapid rate. In both bench and transmission tests, which are run for a fixed period of time, this phenomenon could easily result in misleading conclusions about the relative quality of an ATF. An excellent example of this is the CRC reference fluid, Fluid C in Fig. 3. The actual break point of between 99 and 123 hr in the Enjay bench oxidation test characterizes the fluid as being borderline quality in the Powerglide L-39 test. In other words, it is a fluid which has a break point close to 300 hr in the transmission test. This makes it ideal as a reference oil if time to break point is measured. But if the transmission results are analyzed only after 300 test hr, it might well be a poor reference oil. This is because the normal test variability would probably result in transmissions ranging from very clean to very dirty depending on how close the break point was to 300 hr. Other factors such as transmission test setup and operating controls would further complicate the problem of trying to standardize the 300-hr results with this fluid.

These studies also point out the importance of regular changes of automatic transmission fluids to the automobile owner. It is easy to see how a fluid might be quite satisfactory in its performance up to the recommended drain period, but after another few thousand miles could be badly degraded and cause difficulties if left in the transmission.

Shear Stability

Present-day automatic transmissions are designed to operate with fluids having good viscosity characteristics at both low and high temperatures. Good low-temperature fluidity is necessary to prevent transmission damage during cold weather start-up and operation. High viscosities at elevated temperatures are desirable to maintain adequate lubrication, reduce internal leakage, and insure proper shifting characteristics of the transmission. V.I. improvers are commonly used to impart these desirable viscosity characteristics. However, as is well known, V.I. improvers break down under the shearing conditions present in transmissions. This results in a loss in the high-temperature viscosity, which, if too severe, adversely affects transmission operation. Therefore, as with oxidation stability, shear stability of a fluid could be a factor in determining the recommended drain period of the fluid. The importance of shear stability to transmission manufacturers is quite evident, since it is one of the properties measured in the 50-hr Hydra-Matic durability test required for Type A-Suffix A approval of an ATF. We have carried out a field study on the factors affecting shear stability of ATF's. This study

[4] Letter on "Proposed Automatic Transmission Fluid Specifications," from Norman A. Hunstad, secretary, General Motors Committee on Automatic Transmission Fluid, Jan. 14, 1959.

is discussed below and the results are related to the 50-hr Hydra-Matic durability test.

Field Test — A multiunit Mileage Accumulation Dynamometer, commonly referred to as the "MAD" unit,[5] was used to study important factors affecting the viscosity loss of transmission fluids. Briefly, the unit is composed of eight outdoor dynamometer stands in which automobiles equipped with automatic transmissions can be operated under identical driving conditions. Signals from magnetic tapes programmed under actual driving conditions are used to control the operation of the cars. This unit has been valuable for evaluating fluids, because a variety of driving conditions can be exactly duplicated from car to car, thereby permitting an independent evaluation of the other factors affecting the shear stability of transmission fluids. The following three factors have been investigated:

1. Effect of driving conditions.
2. Effect of transmission make.
3. Effect of V.I. improvers.

The fluids used in this study were Type A Suffix A quality. They were formulated with detergent inhibitors imparting excellent oxidation stability in order to minimize viscosity increase due to oxidation. The detailed procedure used in carrying out these studies is described in Appendix II.

Effect of Driving Conditions — The MAD unit was used to compare the effects of high-speed driving versus stop-and-go city driving. As illustrated in Fig. 9, high-speed driving is much more severe on AT fluid shear stability. In this comparison using transmission W and Fluid K, 5000 miles of 70% high-speed, and 30% city driving was equivalent to 20,000 miles of city driving.

For the above comparison, as well as subsequent ones, the V.I. improver breakdown is calculated based on the loss in viscosity at 210 F of the finished fluid divided by the total viscosity contributed by the V.I. improver to the fresh fluid.

Effect of Transmission Make — The make of transmission was found to have an important effect on the shear stability of an ATF. Tests were run on four makes of transmissions and two ATF's under both 70% high speed-30% city driving and 100% city driving conditions. Fig. 10 compares Transmissions W and X operated on the high-speed tape using Fluid J. The transmissions were equally severe up to 5000 miles. Beyond this mileage, Transmission X continued to break down the V.I. improver rather severely, whereas Transmission W tapered off in its severity. These results were duplicated with Fluid K in two different cars having the same makes of transmissions, as is illustrated in Fig. 11.

Although only two comparisons of Transmissions W and X were made, the excellent agreement between Figs. 10 and 11 indicates that the difference in the shear characteristics of the transmissions is due more to design factors than individual transmission variability.

A final comparison may be made between three different makes of transmissions operated on Fluid K under 100% city driving conditions. As shown in Fig. 12, there is a wide range of severity of transmission on V.I. improver breakdown. Transmission Y was the most severe. After less than 5000 miles

[5] "Eight Lane Dynamometer Highway," by O. G. Lewis, R. R. Risher, Jr., and J. A. Wilson. Paper presented at SAE Summer Meeting, Atlantic City, June, 1959.

Table 5 — Repeatability Data on Enjay Bench Oxidation Test

Fluid Identification	1	2	3
Hours to Break Point	288	216	192
	264	216	216
	> 300	192	192
	> 300		192
	284		
	292		
Average	288	208	198

Standard Deviation, $S_p = 13$ hr

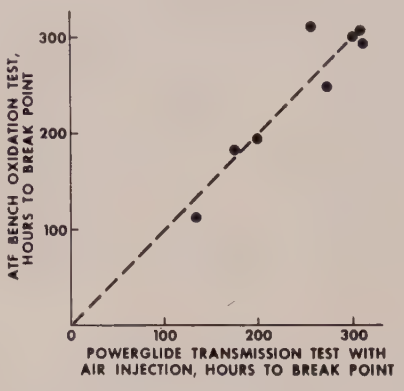


Fig. 8 — Correlation of Enjay Bench Oxidation Test with Powerglide Test with air injection

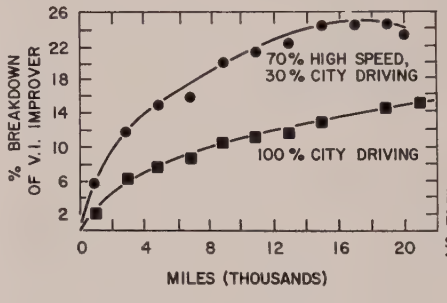


Fig. 9 — Effect of driving conditions on ATF shear stability using Transmission W and Fluid K

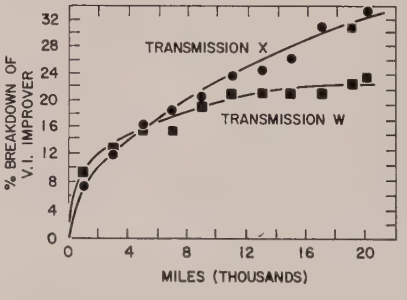


Fig. 10 — Effect of transmissions on ATF shear stability using Fluid J and 70% high speed—30% city driving

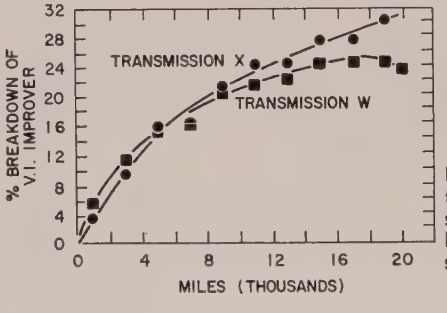


Fig. 11 — Effect of transmissions on ATF shear stability using Fluid K and 70% high speed — 30% city driving

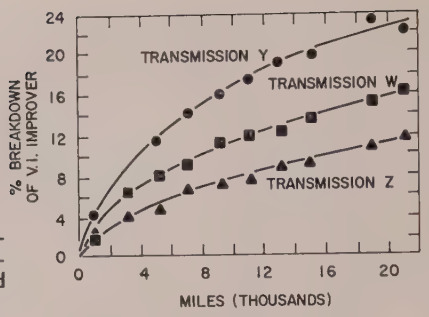


Fig. 12 — Effect of transmissions on ATF shear stability using Fluid K and 100% city driving

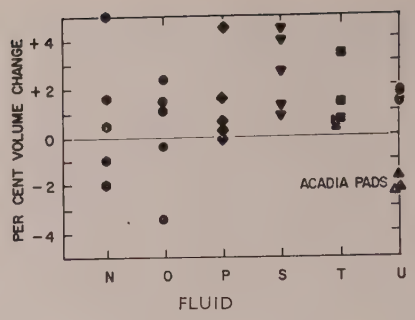


Fig. 14 — Rubber swell measurements on commercial automatic transmission seals show poor reproducibility

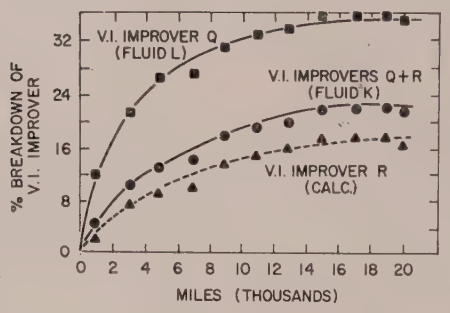


Fig. 13 — Effect of V.I. improvers on ATF shear stability using Transmission W and 70% high speed — 30% city driving

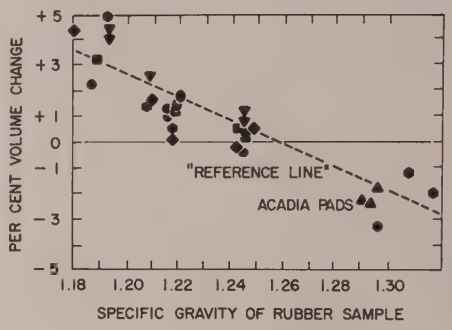


Fig. 15 — Replot of data in Fig. 14 shows rubber swell is function of specific gravity of transmission seal

of city driving, it sheared the fluid as much as Transmission W did in 10,000 miles and Transmission Z did in 20,000 miles.

In fact, by comparing Figs. 9 and 12, it may be seen that 100% city driving with Transmission Y is almost as severe as 70% high speed-30% city driving with Transmission W.

Effect of V.I. Improvers — A third factor affecting the shear stability of ATF's is the concentration and inherent shear stability of the V.I. improver in the fluid. Fig. 13 illustrates MAD results on two fluids similar in every respect except in the V.I. improver treatment. Fluid L containing V.I. Improver Q exhibits a breakdown of 35% after 20,000 miles in Transmission W. Fluid K containing a combination of V.I. Improvers Q and R loses only 20% of the viscosity increase due to the V.I. improvers. From these data, a breakdown of 16% was calculated for V.I. Improver R.

The Type A Suffix A specification requires that AT fluids must not break down to a viscosity less than 46.5 SUS at 210 F in the 50-hr Hydra-Matic durability test. A formulator of an ATF must be able to predict the breakdown of his fluid to assure its meeting this specification. This may be done in either of two ways. First, a laboratory bench test such as a sonic oscillator may be correlated with the durability test and used for the prediction. Second, breakdown values for the V.I. improvers may be used to calculate the viscosity loss. The values of 35% for V.I. Improver Q and 16% for V.I. Improver R have proved useful for this purpose. A comparison of actual and predicted viscosity losses of two fluids employing mixtures of V.I. Improvers Q and R is shown in Table 6.

The breakdown values of 35% and 16% used in these calculations were selected on the basis of the relationship of the 50-hr durability test to the 20,000-mile field results.

Relationship of 50-hr Hydra-Matic Durability Test to Field Results — Two of the fluids run on the MAD unit have also been run in the 50-hr Hydra-Matic durability test. A comparison of the results is shown in Table 6.

Results obtained on the two fluids are in good agreement indicating that the 50-hr durability test is equivalent in shearing action to 10,000 miles in Transmission X or 20,000 miles in Transmission W, both operated on the 70% high speed-30% city driving schedule. From this comparison, an estimate may be made of the relationship of the 50-hr durability test to other transmissions and other driving conditions. This has been done as shown in Table 7.

According to a survey made by the U. S. Bureau of Public Roads, 54% of the mileage accumulated in the U. S. is on rural highways and 46% is on city streets.[6] Thus, the average motorists' driving habits are bracketed by the two tapes used in the MAD units studies above. On this basis, it may be concluded that the 50-hr Hydra-Matic durability test gives a reasonable estimate of the viscosity loss occurring in an ATF within the recommended drain period.

Rubber Swell Measurements

The problems of measuring rubber swell characteristics of AT fluids have plagued lubricant refiners for some time. Variables which affect the test results are the base oil type, additive type, and rubber composition. Bozzelli[7] discussed the poor reproducibility of the standard ASTM D-471-57T test even when "standard" rubber samples were used. Kite and Koenig[8] expressed the opinion that the ASTM test should not be used as a measure of performance. They felt that transmission oxidation or

[6] Table VM-I, 1956. U. S. Bureau of Public Roads.
[7] "Designing Automatic Transmission Fluid," by A. J. Bozzelli. Paper presented at SAE Annual Meeting, Detroit, January, 1959.
[8] "Evaluating Automatic Transmission Fluids," by W. H. Kite, Jr. SAE Special Publication 161, 1959.

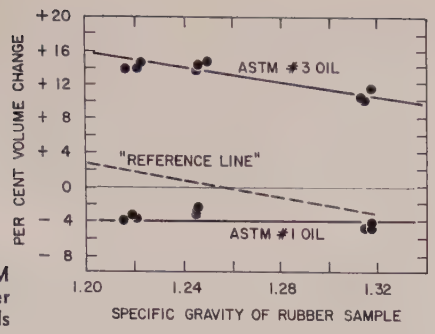


Fig. 16 — Effect of ASTM Reference Oils on rubber swell of commercial seals

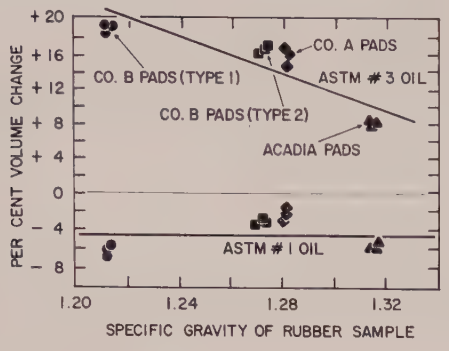


Fig. 17 — Effect of ASTM Reference Oils on rubber swell of standard rubber compounds

Table 6 — Viscosity Loss in 50-Hr Hydra-Matic Durability Test versus MAD Field Test

	Viscosity Loss, SUS at 210 F	
	Fluid J	Fluid K
50-Hr Hydra-Matic Durability Test		
Predicted[a]	2.6	2.2
Actual	2.6	1.9
MAD Field Test		
Transmission X, 10,000 miles[b]	2.4	2.0
Transmission Y, 20,000 miles[b]	2.6	2.2

[a] Using breakdown values of 35% for V.I. Improver Z and 16% for V.I. Improver R.

[b] 70% high speed — 30% city driving.

Table 7 — Relationship of 50-Hr Hydra-Matic Durability Test to Field Performance

Transmission	Recommended Drain Period, Miles	Mileage Required to Give Viscosity Loss Equivalent to 50-hr Hydra-Matic Durability Test	
		100% city driving	70% high-speed 30% city driving
W (1957)	25,000	40,000 (est.)	20,000
X (1958)	12,000	—	10,000
Y (1958)	20,000	22,000	—
Z (1959)	16,000	60,000 (est.)	—

durability cycling tests gave more reliable information on seal performance. However, even in these tests they found that the commercial seals gave poor rubber swell reproducibility. Studies carried out by our laboratories indicate that the poor reproducibility is due primarily to variations in the rubber seals. A technique has been developed for taking the variability into account, thus permitting a reasonable interpretation of otherwise useless data.

We obtained a batch of the standard rubber pads, Acadia 3612, as specified in the Type A, Suffix A specification for rubber swell studies. These pads were soon found to give shrinkage with most commercial ATF's. Therefore, a switch was made to the use of commercial seals for rubber swell measurements. Fig. 14 shows the data obtained with these commercial seals on six different automatic transmission fluids. It may be seen that the results vary quite widely. For example, Fluid N varied from −2 to +5 volume % change. Careful examination of the data showed a wide variation in the specific gravity of the seals. This indicated that either the compounds used in manufacturing the seals were quite different, or the ratio of polymer to filler was different. Fig. 15 shows the same data plotted as a function of the specific gravity of the rubber samples. This plot demonstrates that a considerable tightening in the rubber swell data can be obtained by taking specific gravity into account. As shown in Fig. 15, the specific gravities of the rubber seals varied widely from as low as 1.18 to above 1.31. Further, as the specific gravity of the seal increased, the volume swell decreased. The tendency for the Acadia pads always to give shrinkage appears quite apparent on this plot, since the pads fall in the very high specific gravity range.

Fig. 16 shows the volume change as a function of specific gravity of the rubber sample when tested with the standard ASTM oils. ASTM No. 3 oil, a highly naphthenic oil, gives high-volume swell with all specific gravity groups. However, again the amount of swell appears to be a function of the specific gravity of the rubber sample. With ASTM No. 1 oil, a highly paraffinic oil that gives volume shrinkage, the specific gravity of the rubber sample does not appear to affect the per cent volume change. For reference purposes, a dashed line showing the reference line from Fig. 15 is included.

Additional studies have been carried out with the ASTM reference oils and test pads of various transmission manufacturers. Four different standard pads supplied by three different automotive companies were used. In Fig. 17 it can be seen that in general the specific gravity of the rubber sample is an important factor in determining the amount of swell obtained with ASTM No. 3 oil. This is apparently true even though the pads themselves may be of entirely different compositions. Again with ASTM No. 1 oil, the volume change does not appear to be affected by the specific gravity of the rubber sample.

The reason for the importance of specific gravity of rubber in determining volume swell has not been thoroughly investigated. One possibility is that it is a measure of the amount of polymer in the compound. It is expected that compounds having high concentrations of polymer, and thus low specific gravity, would give greater volume swell. There are undoubtedly other factors which should be taken into account in measuring volume swell. The specific gravity factor, however, is an easily measured one which can be used to improve greatly the accuracy and interpretation of rubber swell measurements.

Conclusions

Oxidation Stability —

1. The Enjay Bench Oxidation Test shows an excellent correlation with the Powerglide transmission

test with air injection in measuring the oxidation stability of automatic transmission fluids.

2. Most automaic transmission fluids show a sharp break with time in oxidation stability. Therefore, in either bench or transmission tests designed to measure oxidation stability, the time to break point should be given careful consideration.

Shear Stability —

1. High-speed driving is more severe than city driving in terms of shear stability of an ATF.

2. The type of transmission is an important variable on the shear stability of ATF's.

3. The amount and type of V.I. improvers used in an ATF have a very important effect on shear stability.

4. In general, the 50-hr Hydra-Matic durability test gives a realistic appraisal of the shear breakdown of an ATF encountered in recommended field usage.

Rubber Swell

1. The reproducibility and thus the interpretation of rubber swell measurements may be greatly improved by taking into account the specific gravity of the rubber sample.

2. Volume swell is a function of the specific gravity of rubber seals when fluids containing naphthenic components are being tested.

3. With fluids containing all paraffinic base oils, there appears to be no relationship between rubber swell and the specific gravity of the rubber sample.

APPENDIX I

Enjay Bench Oxidation Test

Objectives — The Enjay bench oxidation test has been designed to study the oxidation stability of automatic transmission fluids. In this test, fluids exposed to bronze catalyst are mechanically agitated in a 300 F oven for a time period ranging from 72 to 300 hr. The fluid's oxidation stability is then evaluated by determining the hours required to form sludge in the fluid as determined by a filter paper test.

Equipment —

1. Oven — forced draft, thermostatically controlled to $\pm$ 1.0 F.

2. Stirring mechanism — glass paddle stirrers operated at 240 rpm.

3. Sampling device — a hypodermic needle attached to a 30-ml syringe is inserted through an opening in the oven closed on the inside by an inverted oil spring cap.

4. Catalyst — one quarter section of a bronze thrust washer from a 1955 Powerglide transmission (part number 2-3697501). The catalyst is sand blasted to give a uniform surface.

Test Procedure —

1. Weigh out 325 grams of fluid in a 600-ml beaker.

2. Suspend catalyst (one-quarter section of a sand-blasted bronze thrust washer) in the fluid by means of a glass hook attached to the beaker. Catalyst should be totally immersed.

3. Insert beakers in oven heated to 300 F and adjust stirrer speed to 240 rpm. The test is timed from this point.

4. Every 24 hr, after the first 72 hr have elapsed, withdraw a portion of the sample with the syringe and needle and vacuum filter 15 drops through Whatman No. 1 paper trimmed to the size of a nickel. Rate the filter paper for sludge deposits using a demerit scale where 0 represents a filter paper completely free of sludge particles and 10 represents a fluid that is unfilterable. Return the unused portion of the fluid to the test beaker.

5. Cease the test when the fluid becomes unfilterable or 300 hr have elapsed.

Note: Other variables such as viscosity increase, acid number increase, and analysis of the used oil for insolubles have been investigated. It was found that a plot of visvosity versus time will define break point of the fluid just as the filter paper test does. The other analyses proved useless for correlation studies. However, for special studies these analyses may be desirable. Therefore, the following inspections are optional:

Fluid

1. Viscosity at 100 and 210 F, initially, every 24 hr and at end of test. Fluid removed from the beaker during test must be returned.
2. Total base number and acid number initially and at end of test.
3. Insolubles at end of test.
4. Evaporation loss (tared beaker required).

Equipment

1. Sludge rating on the drained beaker.
2. Varnish rating on stirrers.

APPENDIX II

ATF Shear Stability Studies Performed on MAD Unit

Automobiles Used — Late model (1957–1959) automobiles from four major automobile manufacturers were used to study the important variables affecting the viscosity loss of an AT fluid in transmission operation. These automobiles had mileages of 700–25,000 at the beginning of the test.

Transmission Flush Procedure — Fluid in the transmission is drained from the sump and the converter or torus. Experimental fluid is added and the transmission operated in the drive position for 20 min to heat the fluid and insure thorough mixing. The transmission is then drained as before and fresh fluid added as a final flush. The transmission is again operated for 20 min in the drive position, the fluid drained and fresh fluid added for testing purposes. This flush procedure removes 97% of the fluid initially in the transmission.

Test Procedure —

1. Automobiles are placed on the MAD unit and operated on one of two tapes.

a. 100% City Driving Tape — Tape was developed from actual driving conditions experienced in Newark, N. J., and its environs. Average automobile speed for this tape is 28 mph.

b. 70% High Speed-30%City Driving Tape — This tape was derived from a combination of high-speed, parkway driving (60 mph) and city driving (20 mph). Seventy per cent of the miles were accumulated on the Garden State Parkway and 30% on the streets of Paterson, N. J.

2. A 2-oz sample of fluid is taken from the transmission at 0, 50, and 1000 miles and every 2000 miles thereafter. Viscosities at 100 and 210 F are determined on these samples.

Discussion of this paper will be found on pp. 465-467, following the Selby paper.

Automatic Transmission Fluid Viscosity at Low Temperature

and its effect on transmission performance

Theodore W. Selby
Research Laboratories, General Motors Corp.

A LOW-TEMPERATURE study of the relationship between the performance of a step-type automatic transmission and the transmission fluid viscosity is reported in this paper.

It is shown that the low-temperature malfunction of these units is due to the viscometric properties of the fluid and that at the temperature at which the fluid reaches a certain critical viscosity the transmission will fail. A mathematical analysis of the mechanism of failure supports the conclusions drawn from the experimental study.*

THE AUTOMATIC TRANSMISSION is a well-known and widely publicized part of today's automobile. Less known are the problems which this device has presented to the automotive engineer. Since the automatic transmission is greatly dependent on the physical properties of the transmission fluid, it is in this area of fluid properties that many problems have arisen.

One of the more important physical properties of the fluid is the coefficient of viscosity. The marked temperature dependence of this property has led to a number of problems in transmission behavior at temperature extremes. Of these, the effect of the viscosity of transmission fluids on low-temperature transmission performance has been particularly troublesome, especially in regard to field experience with a particular make of a step-type transmission. A field problem involving the malfunction of these transmissions during very cold weather initiated the study reported in this paper.

Field Problem

Automatic transmissions are divided roughly into three groups:

1. The step-type transmissions which change the transmission input-output torque ratio in definite steps or shifts (usually three or four) in a manner somewhat similar to that of the manual transmission, except that the shifts are smooth and are made by means of a clutch and gear mechanism while engine torque is being applied.

2. The torque-converter transmissions which change transmission input-output torque ratio continuously by means of a regenerative turbine fluid-coupling.

3. The stepped torque-converter transmissions which are a mechanical combination of the first two, having a regenerative torque-multipling turbine plus a clutch and gear mechanism requiring one or two steps.

During the winters of 1950 and 1951 the importance of good low-temperature fluidity was forcibly brought to the attention of the transmission engineers. The difficulty, which was observed with step-type transmissions during this period, was characterized by a burned condition of the cork and paper clutch plate surfaces similar in appearance to that shown in Fig. 1.

At the time, it was suspected that such malfunction was associated with the low-temperature vis-

* Paper presented at SAE National Fuels and Lubricants Meeting, Chicago, Oct. 29, 1959.

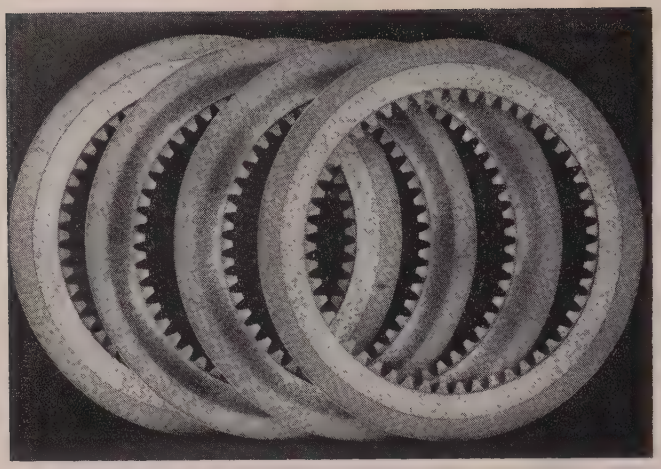


Fig. 1 — Front composition clutch plates from cold-room test

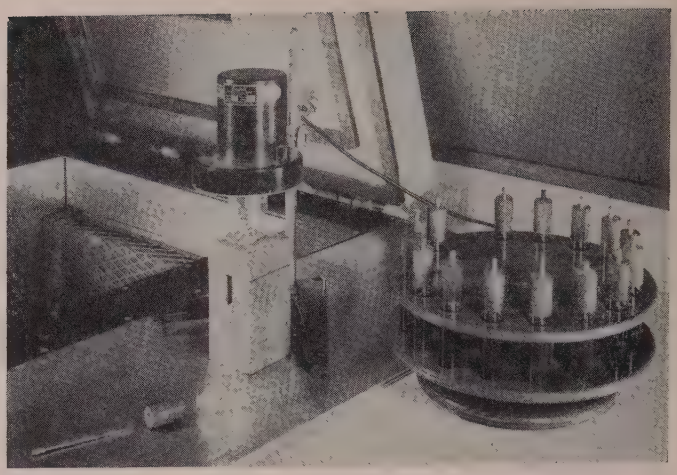


Fig. 2 — Brookfield viscometer and supporting equipment

cous properties of the fluid, but studies of the fluids' pumpability characteristics by other investigators did not show the expected correlation. In 1955 a fresh attack was made on the problem and these studies, which have been continued up to the present time, show a strong relationship of transmission failure and fluid viscosity at low temperatures.

Experimental Efforts

Approach — The problem was divided into three phases:

1. To establish a repeatable full-scale cold-room test which would simulate field failure.
2. To devise a viscometric test by which one could characterize the viscosity-temperature relationship of automatic transmission fluids at low temperatures.
3. To relate the viscometric observations to those of the cold room.

Cold-Room Test Technique — The cold-room experience gained by GM's Oldsmobile Division was very helpful in setting up a repeatable test method. After a fairly extensive series of cold-room tests, during which the condition of the clutch plates was used as the criterion, the following procedure was established:

1. A transmission is disassembled, cleaned in naphtha, and dried.
2. New clutch plates are soaked for approximately 20 min in the fluid under test, after which they are used in the assembly of the clean transmission and the transmission installed in the test car.
3. The transmission is filled with the fluid under test and the car taken out for a 2-mile test run to check the functioning of the engine and transmission. (A pressure gage connected to the main gallery from the transmission's front pump helps in the evaluation of transmission performance.) 5W engine oil is used for lubrication and 14–15 lb Reid vapor pressure gasoline is used for fuel to assist the engine in low-temperature starting.
4. The car is then placed on the cold-room inertia-weight dynamometer whose inertia has been previously adjusted so that the time for a 0–70 mph acceleration for the car on the rolls is the same as on the test track.
5. Thermocouples are put into the transmission sump, the engine coolant, and the torus member. These are connected to a recorder.
6. The rear wheels are jacked up off the dynamometer rolls (to prevent deformation of the tires and the resultant tire bump during the first few accelerations). A fan is aimed to blow on the transmission through a grating in the cold-room floor beneath the car.
7. Cool-down is started at 4:00 p.m. and by 12:00 p.m. to 1:00 a.m. the transmission sump temperature is usually at cold-room temperature. Thus, approximately 8-hr "soak" at the test temperature is obtained.
8. At 8:00 a.m. the thermocouple readings are compared. If there is more than 2 deg difference in any two of them the test is delayed until they are brought into agreement.
9. Just before the first start at 9:00 a.m., the rear wheels are lowered onto the rolls and the torus thermocouple removed. The fluid level in the transmission is also checked and an engine tachometer attached.
10. The car is started using a rectifier to supply the current for the starting motor (batteries were found to be too variable in performance). The engine is allowed to idle at 800 rpm until the transmission pressure gage shows that pump pressure is established. (Some care has to be taken at this point to prevent the engine from stalling as the transmission "comes to life."
11. Immediately the car is accelerated at a manifold vacuum of 7.5 in. of Hg to a speed of 60–70 mph and then quickly braked to a full stop, after which it is immediately reaccelerated and again braked. This test technique continues until a total of six accelerations have been made.
12. The engine is then turned off and the car prepared for another start at the test temperature at 3:30 p.m., at which time another six accelerations are made.
13. The car is again prepared for an overnight cool-down and soak and at 9:00 the following morning the third and final start is made.
14. This series of three starts and 18 accelerations constitutes a test.

15. The car is then removed from the cold room, the transmission disassembled, and all clutch plates in the front and rear packs are closely inspected.

16. Any trace of burning on any one of the clutch plates is sufficient to consider the transmission to have failed.

It may be appreciated that the above test is rigorous and that repeatability would require that the transmissions (three were used) and engine be maintained in good operating condition, that the driver of the car be systematic in his driving procedure, and that the cold-room crew closely control the environment of the test. Even with the best of efforts only fair test uniformity was expected. However, the efforts of the cold-room crew produced a surprisingly high degree of repeatability.

Viscometric Analysis Technique — After repeatable cold-room results were obtained, the next step was to find if these results could be correlated with the viscometric characteristics of the transmission fluids. The marked change in the viscosities of these fluids with temperature (for example, a change from 1000 to 100,000 centipoise (cp) over a temperature range of 0 to – 50 F) required the use of a viscometer capable of responding to a wide range of viscosities.

It was desirable that the method of viscometric analysis be as simple and rapid as possible without the sacrifice of either precision or accuracy, although it was anticipated that involved analytical techniques might be required because of the non-Newtonian characteristics of these fluids at the lower temperatures.

One wide-range instrument which was available to the author was the rotational Brookfield viscometer, Model LVT. Its range of response varied from 10 to 2,000,000 cp with an accuracy claimed to be within $\pm 1\%$ and a precision within $\pm 0.2\%$. However, this viscometer had no integral means of sample temperature control. Moreover, using a No. 4 spindle, it was a relatively low shear-rate device (~ 1.6 sec^{-1} at 60 rpm) and could not be expected to view non-Newtonian (or shear dependent) fluids in the same manner as the transmission which operates at shear rates calculated to be in the order of at least 10,000 sec^{-1}. (See Appendix.) Nevertheless, it was felt that the viscometer would prove to be of value in the region of Newtonian and slightly non-Newtonian behavior, which for automatic transmission fluids may extend up to 10,000 cp.

Fig. 2 shows the equipment used in the viscometric analysis. To the left is the Brookfield viscometer and below it is a fluid sample of the process of being analyzed. The sample is contained in an analysis cell, several of which are shown in a rotating rack and another shown in the foreground in a disassembled state. The cell under the viscometer is enclosed by a balsa wood carrier which serves as an insulator during the few minutes that the sample is outside of the constant-temperature box prior to viscometric analysis. Double pane plastic windows on opposite sides of the balsa block permit the analyst to position the necked portion of the spindle at the liquid-air interface accurately.

The rotating rack shown in Fig. 2 is used as the cell holder in the constant-temperature box. This rack turns at approximately 32 rpm under the force of the air circulated inside the constant-temperature box. The rotation of the rack insures that each sample analyzed in a given group will be exposed to nearly identical cooling conditions. This was found to be particularly important in early experiments with fluids showing marked non-Newtonian behavior.

The constant-temperature box, partially shown in Fig. 2, was capable of maintaining temperature control of ± 0.3 F at any temperature from 0 to – 50 F and was equipped with a double-compressor, Freon 13–22 cascade system to produce the temperatures desired. The analytical procedure developed was as follows:

Twenty-five milliliters of each fluid to be analyzed was poured into individual cells (the glass container of which was a 25×100 mm test tube). The cells were then placed in the rack and the rack put into a 125 F oven for a half-hour (to assure, by increasing the motion of the oil molecules, that any predisposition to ordered structures or "memory" of the fluids was removed without heating, such "memory" was found to influence the fluids' low-temperature viscosity values). The rack, cells, a mercury-thallium thermometer, and the balsa carrier were then put into the constant-temperature box and the box set at the desired temperature.

The next morning, after the box temperature had been read from the thermometer, one cell at a time was transferred from the box to the viscometer via the balsa carrier. Care was taken to keep the box lid closed as much as possible to maintain the desired temperature in the remaining unanalyzed samples.

When all of the samples had been analyzed, the rack was again placed in the oven at 125 F for a half-hour. In the afternoon the rack was put back into the constant-temperature box for analysis at a new temperature the following morning. (A modified procedure of using new samples of the fluid for each temperature gave the same results as re-analysis of the same sample.)

Initially, viscometric analyses were made at 5 F intervals from 0 to – 40 F, but later work showed that the viscosity-temperature curves constructed using these values could be essentially duplicated by analyzing fluids at 10 F intervals and that this latter technique would make it possible to complete the whole series of temperatures in a week. To complete the five analyses (0, – 10, – 20, – 30, and – 40 F) with four overnight cool-down and soak periods, the 0 F cool-down and soak was run during one of the working days with the samples put into the box in the morning and analyzed 6 hr later. This shorter soak period for the 0 F analysis was considered valid since the fluids were Newtonian at this temperature and thus only the cool-down period of 3–5 hr was really required before analysis.

The repeatability of this viscometric technique was indicated to be at least within $\pm 10\%$ standard deviation over the temperature range of 0 to – 40 F. This was considered reasonably good when it was observed that the viscosity change with temperature may exceed 10% per degree Fahrenheit at lower temperatures.

The obvious sources of error were: (1) temperature control while transferring samples from the box to the viscometer, and (2) accuracy of the thermometer both intrinsically and in regard to its placement in the box relative to the position of the rack. In regard to the intrinsic thermometry error, this was reduced, by using mercury-thallium ther-

mometers instead of pentane- or alcohol-filled thermometers.

Cold-Room Test Results — The cold-room tests on a fluid were run by decreasing the temperature in 5 F steps until failure was observed. Thirty-four fluids were analyzed in this manner. Most of these fluids qualified under the Automatic Transmission Fluid, Type A program, but several were experimental fluids and one was a nonqualified fluid which appeared on the market.

It was found early in the course of the tests that the severity of the clutch plate burns became progressively greater the farther the test temperature was below the failure temperature. This is shown in Fig. 3 where a series of clutch plates ranging in burn intensity from extremely light (second across from the upper left corner) to extremely heavy (at the bottom) are contrasted. (An unburned plate is shown for comparison in the upper left corner.) The progressive severity of burning shown was found to represent a decrease in 1 F steps from the incipient failure temperature denoted by the clutch plate with the extremely light burn. Using this scale, a moderate burn obtained at a given test temperature could be interpreted to indicate that the temperature of incipient failure was 4 F higher. Thus, a more precise comparison of the failure temperatures of the different transmission fluids could be made. It was interesting to note that when light burns were obtained they were almost invariably found in the front clutch plate pack. This observation will be referred to in later discussion.

The repeatability of the cold-room tests was quite satisfactory considering the possible sources of error. It was found that, in general, a given failure could be reproduced within ± 1 F.

The results of the evaluation of 34 fluids in the cold room are given in Table 1. It is evident that the failure temperatures of few fluids lie in the temperature range below – 20 F and fewer still below – 25 F. Most of the failure temperatures lie in the range between – 15 and – 20 F. One failure temperature, that of the unqualified fluid, lies above – 10 F.

When the above investigation was about two-thirds completed, it was noticed that repeat analyses of the same fluids gave significantly lower failure temperatures. It was thought that this loss of repeatability could have been caused by loss of engine power due to the long and severe low-temperature service to which it was subjected. The engine was inspected, found to be quite worn and replaced. Subsequent repeat tests at low-temperatures duplicated earlier work.

Brookfield Viscometer Test Results — The 34 fluids tested in the cold-room were analyzed with the Brookfield viscometer. The viscosity values of each fluid obtained at the five temperatures were plotted versus temperature on semi-logarithmic paper and a smooth curve drawn through the points. Several of these curves are shown in Fig. 4. In Table 2 are shown the viscosity values taken from the curves at the nominal temperatures of 0, – 10, – 20, – 30, and – 40 F, plus the viscosity values at the failure temperatures recorded in Table 1.

Correlation of Cold Room and Viscometric Results

Relative Constancy of Failure Viscosity — It is evident from Table 2, that transmission fluids may vary markedly in their viscosity-temperature relationships. Fluid O, as an extreme example, passes the cold-room test at – 10 F yet becomes virtually solid below – 25 F (for which viscous behavior it was disqualified). However, in spite of the obvious variation of viscosity with temperature for each fluid and among the fluids at a given temperature, the failure temperatures for this series of fluids are all fairly similar — with an average of 4910 cp and a standard deviation of ± 1200 cp. It should be noted that this deviation includes the systematic errors of both the cold-room test and the viscometric analyses.

Variation of Failure Viscosity with Test Temperature — If the data given in Table 2 are more closely inspected, it appears that the failure viscosity increases somewhat with decreasing temperature. This is shown in Fig. 5 wherein the failure viscosities are plotted versus temperatures. The circles around each point are meant to convey a conception of the magnitude of error. This "area of error" reflects the ± 1 F variation of the cold room about the recorded failure temperature and the error in precisely judging the incipient failure temperature from the condition of the clutch plates as well as the ± 10% standard deviation of the viscosities.

It appears, then, from Fig. 5 that the failure viscosity is roughly doubled (from 3500 to 8000 cp) as the incipient failure temperature drops from – 10 to – 30 F.

Discussion

Mechanism of Clutch Plate Burning — It is quite apparent that the burning of the cork-paper clutch surfaces require high surface temperatures. Ex-

Table 1 — Cold-Room Evaluation of Transmission Fluids

Class	Code Letter	Nominal Test Failure Temperature, F	Interpreted Failure Temperature, F
Experimental	A	– 20	– 17
Qualified	B	– 15	– 15
Qualified	C	– 25	– 24
Qualified	D	– 15	– 13
Qualified	E	– 15	– 13
Qualified	F	– 20	– 17
Qualified	G	– 20	– 19
Qualified	H	– 20	– 18
Qualified	I	– 15	– 15
Qualified	J	– 15	– 14
Qualified	L	– 20	– 20
Disqualified	O	– 15	– 15
Qualified	P	– 30	– 30
Qualified	Q	– 15	– 12
Qualified	R	– 20	– 17
Qualified	S	– 15	– 13
Qualified	T	– 20	– 18
Qualified	V	– 20	– 18
Qualified	W	– 20	– 19
Qualified	X	– 20	– 20
Experimental	Y	– 15	– 13
Experimental	Z	– 15	– 15
Qualified	AA	– 20	– 20
Qualified	AB	– 30	– 26
Qualified	AF	– 25	– 23
Experimental	AG	– 10	– 10
Unqualified	AI	– 10	– 7
Qualified	AJ	– 25	– 20
Experimental	AM	– 20	– 17
Qualified	AN	– 30	– 26
Qualified	AO	– 15	– 11
Qualified	AP	– 20	– 18
Qualified	AQ	– 20	– 18
Experimental	AS	– 15	– 13

periments with a modified soldering iron as a heat source indicated that temperatures in the neighborhood of 450 F are necessary to cause incipient burning. It is difficult to avoid the conclusion that such temperatures can be accounted for only by the heat generated during clutch engagement, as the plates are forced together under high torque while still in relative motion. One can reasonably speculate that if heat is generated at the clutch plate surface faster than it can be removed and if the period of slipping contact is sufficiently long and if the torque under which the clutch plates are engaged is sufficiently high, burning will occur. Such speculation is well supported by the accumulated experience of transmission engineers to whom the time of clutch plate engagement and the torque transmitted through the clutch plate pack are design variables of primary importance.

The time of clutch plate engagement is controlled by the size of an orifice which meters fluid to the clutch apply piston. The dimensions of this orifice are predetermined by the rate of flow of transmission fluid at normal operating temperatures. To a first approximation, this rate of flow is inversely proportional to the fluid viscosity; thus, it is apparent that the time of clutch plate engagement will become longer as the fluid viscosity increases. It is further apparent that ultimately the time of clutch

Table 2 — Viscosity Values

Class	Code Letter	Interpreted Failure Temperature, F	Viscosity at 0 F, cp	Viscosity at −10 F, cp	Viscosity at −20 F, cp	Viscosity at −30 F, cp	Viscosity at −40 F, cp	Failure Viscosity, cp
Experimental	A	−17	1180	2370	5000	11300	27500	3990
Qualified	B	−15	1550	3180	7650	21700	71500	4840
Qualified	C	−24	1100	2050	4150	9300	27000	5650
Qualified	D	−13	1870	3960	8550	18800	51000	4970
Qualified	E	−13	1700	3550	7900	19000	48000	4500
Qualified	F	−17	1070	1970	3940	9400	61000	3150
Qualified	G	−19	1170	2310	4700	9850	22000	4390
Qualified	H	−18	1280	2450	5250	11700	27300	4500
Qualified	I	−15	1600	3030	6150	14600	66000	4250
Qualified	J	−14	1740	3330	7250	16800	41500	4500
Qualified	L	−20	1750	3460	7200	16100	40000	7200
Disqualified	O	−15	1480	2800	17000	—	—	5100
Qualified	P	−30	840	1760	3730	8250	19900	8250
Qualified	Q	−12	1860	3820	8250	19800	63000	4450
Qualified	R	−17	1250	2330	4600	10300	33500	4000
Qualified	S	−13	1400	2560	4840	10300	27500	3100
Qualified	T	−18	1430	2620	5050	11300	37000	4400
Qualified	V	−18	1300	2780	6500	16600	46000	5450
Qualified	W	−19	1400	2680	5400	12300	37000	5000
Qualified	X	−20	1110	1800	3870	13000	77000	3870
Experimental	Y	−13	1570	3150	6900	16100	39000	3950
Experimental	Z	−15	1490	3060	6650	15300	36300	4500
Qualified	AA	−20	1470	2800	5650	12200	31000	5650
Qualified	AB	−26	1350	2410	4640	10100	35800	7200
Qualified	AF	−23	1260	2370	4750	10900	38500	5950
Experimental	AG	−10	1870	3700	8000	19700	64000	3700
Unqualified	AI	− 7	3350	7900	19900	55500	172800	6050
Qualified	AJ	−20	1580	3280	7000	16100	40500	7000
Experimental	AM	−17	1500	2900	6000	13500	35500	4750
Qualified	AN	−26	1210	2180	4150	8300	18000	6250
Qualified	AO	−11	1800	3620	7750	18300	50000	3890
Qualified	AP	−18	1210	2420	5240	12300	30500	4450
Qualified	AQ	−18	1490	2460	5250	15500	61000	4400
Experimental	AS	−13	1590	3090	6300	14500	47000	3800
							Average	4910
							Standard Deviation	±1200

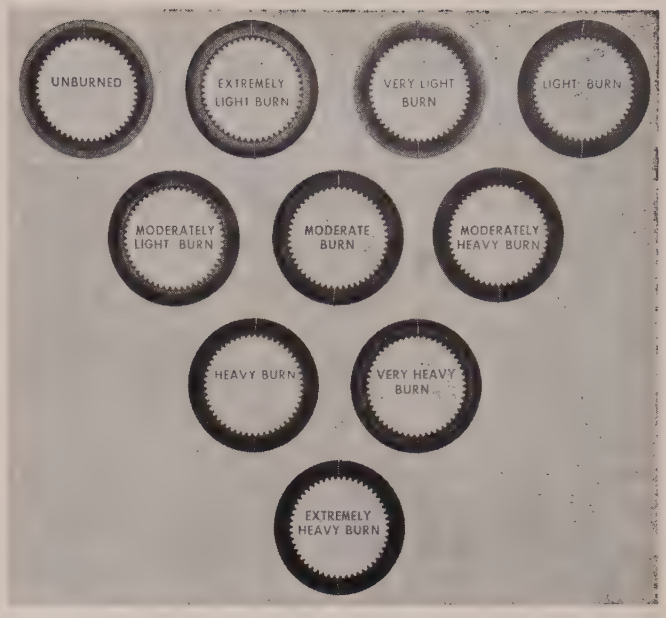


Fig. 3 — Progressively burned clutch plates from various cold-room tests

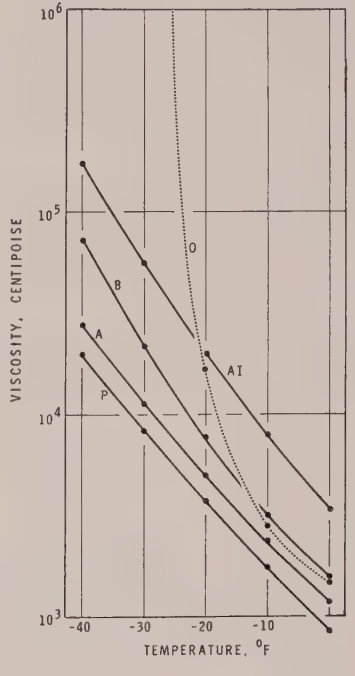


Fig. 4 — Viscosity-temperature characteristics of various automatic transmission fluids from 0 to −40 F

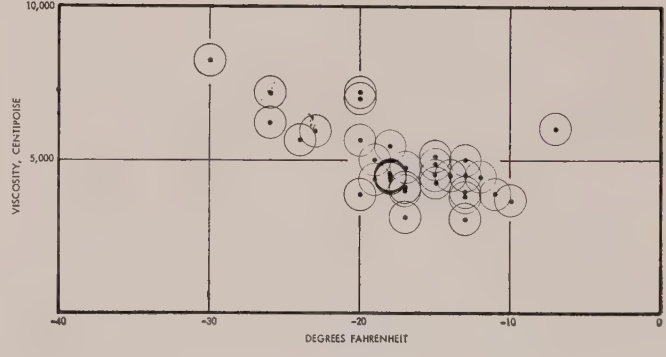


Fig. 5 — Failure viscosity versus failure temperature

engagement will be so prolonged, under a sufficient torque, that the heat due to friction will cause burning of the clutch plates. Moreover, one would expect (other factors remaining constant) that the transmission would fail whenever the viscosity of any fluid used would exceed a given value. In other words, one would expect a failure viscosity and anticipate that it would be independent of temperature.

Calculated Value of Failure Viscosity — In the course of the above experiments and the concomitant efforts to understand the mechanism of clutch plate burning, the author attempted to calculate the critical failure viscosity from certain transmission dimensions as well as from observed operational characteristics of the transmission.

As previously noted, front clutch plate failure seemed to occur preferentially. A satisfactory explanation for this failure pattern was found in the fact that: (1) the front clutch plate pack is the first one engaged after the cool-down and cold soak where the conditions are best for producing burned plates, and (2) the front clutch plate pack is engaged twice (in the 1-2 and 3-4 shifts) during an acceleration through the gears while the rear clutch is engaged only once (in the 2-3 shift).

(Some support for this explanation is found in the record of failures for an earlier model of this transmission which employed a second gear start. Many of the field and cold-room failures with this earlier transmission involved burning of the rear clutch pack as would be expected from the above analysis of the mechanism of clutch plate burning.)

The above considerations served to focus attention on the front clutch operation. The transmission dimensions required for calculation were:

1. Those of the orifice controlling the rate of fluid flow into the front clutch apply cylinder and the accumulator.
2. The volume of fluid necessary to fill the front clutch apply cylinder and the accumulator when the clutch plates were fully engaged.

The other necessary values were:

1. The length of time for clutch engagement under incipient failure conditions.
2. The transmission pump pressure at the temperature of incipient failure.
3. The pressure (due to the clutch drum springs and piston friction) resisting clutch engagement.

The critical dimensions were readily obtainable; however, it was difficult to obtain some of the other values with the desired accuracy. For example, the transmission pump pressure at low temperatures was obtained from the pressure gage used to check out the transmission performance on the test track. At low temperatures the pressure was 100 ± 5 psi. The length of time for clutch engagement under incipient failure conditions was also obtained from this gage by noting how long the gage needle remained deflected during the first shift when pump pressure drops until the clutch apply cylinder and the accumulator are filled. The time of deflection was measured with a stop watch. The critical interval was found to be approximately 4 sec.

The following assumptions were required to calculate the failure viscosity: (1) the rate of flow of viscous fluid through the orifice is governed by the Poiseuille-Hagenbach equation, and (2) the total time of clutch plate engagement (as indicated by the pressure drop) is significant in developing a burned clutch plate. These values and the calculations are given in the Appendix, as well as a critical evaluation of the assumptions.

Notwithstanding the crudity of determining some of the critical values and the limitations of the assumptions, it is still of some interest to find that these calculations yield a failure viscosity of 5300 cp which is in reasonable agreement with the experimental value.

The experimental and mathematical studies reported above partially substantiate the expectation of a constant value of the failure viscosity except that, as noted, a certain temperature dependence of the failure viscosity was found experimentally.

Temperature Dependence of Failure Viscosity — The temperature dependence of the failure viscosity can be explained when one considers *both* of the factors involved in clutch plate burning, that is, time of clutch engagement *and* the engine torque transmitted through the clutch plate pack. While, as mentioned above, it is expected that the failure viscosity would be independent of temperature if the torque (and other factors) remained constant, it is unlikely that in the cold-room tests the torque was constant. Rather it would be expected that the torque would decrease with temperature due to the effect of such a temperature change on the engine oil viscosity.

A 5W engine oil was used. Studies have indicated that the viscosity of such a fluid increases from approximately 2500 at – 10 F to 15,000 cp at – 30 F (this is the temperature range through which the value of the failure viscosity changed from about 3500 to 8000 cp). Such a viscosity change in the engine oil would certainly be expected to provide increased engine drag due to viscous traction.

The argument that decreased torque output of the engine can lead to an increase in the value of the failure viscosity is supported by the aforementioned incident in the cold-room study where the engine had to be replaced because of the effect of wear in reducing the torque output.

The fact that the engine oil viscosity changes fairly rapidly once the unit is in operation does not invalidate the above reasoning, since it is believed that the burned condition of the plates occurs in the first or second acceleration of each start. By the time the engine oil warms up enough to permit the engine to operate at full power, the automatic transmission fluid is also warmed and the transmission is more capable of handling the increased load. The initial starting conditions thus decide whether or not the fluid will cause failure; and if initial engine torque is temperature dependent, so too, will be the failure viscosity, at least to some degree.

Perhaps the strongest evidence in favor of the above reasoning was a cold-room test run on an experimental automatic transmission fluid which showed incipient failure at – 61 F. The engine oil was a synthetic high V. I. material whose viscosity was 5000 cp at the failure temperature. The failure viscosity of the fluid was observed to be 3800 cp, in agreement with the transmission failure viscosities obtained when the 5W engine oil viscosity was also approximately 5000 cp.

There are very likely other fluid properties involved in the mechanism of low-temperature clutch plate burning, such as variation in the effect of dif-

ferent transmission fluids on the coefficient of friction between the clutch plates, non-Newtonian viscometric behavior, and the like. It is the author's opinion, however, that the effect of these variables are secondary compared to those of transmission fluid viscosity and engine torque output. Perhaps, the effect of these secondary variables are indicated by the scatter observed in Fig. 5 which seems to be greater than experimental error.

Although all of the cold-room results so far reported in this paper had been obtained with one particular make of step-type transmission, recent tests have shown that similar clutch plate failures can be obtained with a make of the stepped torque converter. In this latter case, the temperature of failure was about 5 F lower than that of the first make of transmission, using any given fluid.

Conclusions

The study of the relationship between the viscous characteristics of automatic transmission fluids and their effect on the low-temperature performance of the transmission reported in this paper was dependent on the development of two tests: (1) a full-scale cold-room test simulating field experience, and (2) a viscometric technique covering a wide viscosity range over the temperature interval of 0 to – 40 F.

The use of these test methods showed that the low-temperature clutch plate failure typical of step-type transmissions can be directly related to the viscosity of the automatic transmission fluid.

In experiments with 34 transmission fluids it has been shown that at a critical viscosity value, called the failure viscosity, the transmission failed according to the standards used in judging the condition of the clutch plates. In comparison to the strong temperature dependency of the viscosity at low temperatures, as a first approximation, the value for the failure viscosity was found to be fairly constant, that is, 4900 ± 1200 cp. A mathematical analysis of the transmission function at low temperatures gave a failure viscosity of approximately 5300 cp which was in reasonable agreement with the experimental value.

Closer inspection of the experimental data showed a certain degree of temperature dependence for the value of the failure viscosity where the failure viscosity increased with decreasing temperature. This temperature dependence was indicated to be related to the effect of engine oil viscosity on the engine output torque.

In view of the fact that automatic transmissions are being used extensively throughout the United States and Canada, it is to be expected that more attention would be directed to improvement of the viscometric characteristics of these fluids both at low temperatures and at high temperatures. Recent efforts on the parts of the automotive and petroleum industries have resulted in significant advances in this direction.

Acknowledgment

The author would like to thank the members of the cold-room team of the Research Laboratories who made every effort to control the numerous variables involved in the full-scale work. The efforts of W. F. Scruggs and T. M. Verdura in analysing the fluid samples were very much appreciated.

It is to each individual's credit that this paper could be written. This is especially true in regard to the active participation and encouragement of N. A. Hunstad throughout the period of these studies.

APPENDIX

Argument

Let it be assumed that the completion of a shift involving the front clutch plates must be accomplished in a given maximum of time to prevent the clutch plates from burning and that the rate of transmission fluid flow through the smallest orifice is governed by the Poiseuille-Hagenbach equation.[1,2]

$$\eta = \frac{\pi P R^4}{8LQ} \tag{1A}$$

where:

η = Coefficient of viscosity
P = Pressure drop across capillary
R = Radius of capillary
L = Length of capillary
Q = Rate of flow

or:

$$\eta = \frac{\pi R^3}{4Q} \cdot \frac{RP}{2L} = \frac{1}{\text{maximum shear rate}} \cdot \text{maximum shear stress} \tag{1B}$$

Now the mean critical shear rate $\bar{G}$ is proportional to the flow rate through the orifice and the mean shear stress $\bar{S}$ is proportional to the pressure drop across the orifice. If one knew the average shear rate existing at the minimum flow rate necessary to prevent burning of the clutch plates, $\bar{G}_{\min}$, and the average shear stress $\bar{S}$ necessary to maintain this shear rate, the critical fluid viscosity η_{cr} could be calculated according to the equation:

$$\eta_{cr} = \frac{\bar{S}}{\bar{G}_{\min}} \tag{2}$$

Critical Values

The critical values for the particular step-type transmission discussed previously at the incipient failure temperatures are:

1. Flow rate controlling orifice dimensions:
   diameter — 0.198 cm (0.078 in.)
   length — 0.292 cm (0.115 in.)
2. Volume of fluid required to fill the front clutch pack cylinder and the accumulator — 51.3 cm^3 (3.13 $in.^3$)
3. Transmission pump pressure — 6.9×10^6 dynes/cm^2 (100 psi)
4. Resisting pressure — 1.72×10^6 dynes/cm^2 (25 psi)
5. Critical time for the 1-2 shift — 4 sec

Calculations

Calculation of Minimum Mean Shear Rate, $\bar{G}_{\min}$ — From the above assumptions 51.3 cu cm of fluid must pass through the orifice in at least 4 sec. If, as assumed, the flow through the orifice is defined by the Poiseuille-Hagenbach equation, then the mean

[1] Article by J. L. M. Poiseuille. Mém. Savants Étrangers, Vol. 9, 1846, p. 433.
[2] Article by Hagenbach. Pogg. Ann., Vol. 109, 1860, p. 385.

shear rate may be calculated by means of the Kroepelin equation:[3]

$$\frac{\overline{dv}}{dr} = \frac{8\overline{v}}{3R} \tag{3}$$

where:

$\frac{\overline{dv}}{dr}$ = Mean shear rate

$\overline{v}$ = Mean fluid velocity

R = Capillary radius

The mean fluid velocity $\overline{v}$ may be calculated from the equation

$$\overline{v} = \frac{Q}{\pi R^2} \tag{4}$$

Thus:

$$\frac{\overline{dv}}{dr} = \frac{8}{3}\frac{Q}{\pi R^3} \tag{5}$$

(This expression for the average shear rate may be contrasted to the expression for the maximum shear rate in Eq. 1B where $\frac{dv}{dr} = \frac{8}{2\pi R^3} = \frac{4Q}{\pi R^3}$)

Substituting the critical values for the symbols:

$$\overline{G}_{\min} = \frac{\overline{dv}}{dr} = \frac{8}{3}\frac{12.82\ \frac{\text{cm}^3}{\text{sec}}}{(0.099)^3\text{cm}^3} = 11100\ \text{sec}^{-1}$$

Calculation of Mean Shear Stress $\overline{S}$ — Minimum transmission pump pressure was measured as approximately 100 psi during cold-room operation. This pressure exerted on one side of the orifice is partially counterbalanced by the clutch pack springs, and the rest, providing a resisting pressure of 25 psi on the opposite side of the orifice. The resultant pressure drop across the orifice is estimated to be about 75 psi or 5.2×10^6 dynes/cm^2.

Now since the mean shear rate is found to be at a distance of 2/3 R from the center of a capillary, then for a Newtonian liquid the mean shear stress must be found at the mean shear rate and:

$$\overline{S} = \frac{RP}{3L} \tag{6}$$

Again substituting the appropriate critical values:

$$\overline{S} = \frac{0.099\ \text{cm} \times 5.2 \times 10^6\ \text{dynes/cm}^2}{3 \times 0.292\ \text{cm}}$$

$$= 0.58 \times 10^6\ \text{dynes/cm}^2$$

Calculation of Critical Viscosity — Substituting the derived values of mean shear rate and stress into Eq. 2 it is found that

$$\eta_{cr} = \frac{\overline{S}}{\overline{G}_{\min}} = \frac{0.58 \times 10^6\ \text{dynes/cm}^2}{1.11 \times 10^4\ \text{sec}^{-1}} = 53\ \text{poise} = 5300\ \text{cp}$$

Discussion

By the above mathematical analysis the author has attempted to calculate the transmission fluid viscosity associated with the critical performance characteristics of a particular step-type transmission at low temperature. Although the calculated value seems to agree rather well with the determined value of 4910 cp, the calculated value must be considered to be only an agreement well within an order of magnitude since the assumptions and relatively crude measurements used during the work allow no greater confidence.

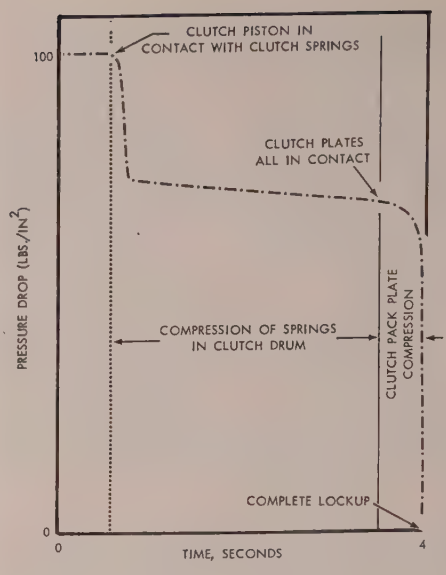


Fig. 6 — Suggested variation of line pressure drop across orifice with time during front clutch engagement

However, neglecting the crudity of some of the physical measurements involved, it may be of some interest to examine the assumptions made. They are reviewed in the order of importance.

Critical Time Interval for 1–2 Shift — In his mathematical analysis, the author assumed that the whole interval of the 1–2 shift (as recorded from the period of needle deflection) was critical in regard to clutch-plate burning. This assumption may be open to some question since it is difficult to picture the occurrence of clutch burning without actual contact of the clutch-plate surfaces under high torque. However, an answer to the question may lie in the fact that in the disengaged position, the clutch plate in the pack closest to its neighboring plate is only at best 0.028 in. distant (that is, in the case where all the clutch plates are equally spaced) and is more likely to be in contact with the neighboring plate. The relative motion of the clutch plates in light contact for a short period of time may produce a polishing of the engaging surfaces. This would result in at least two surfaces which, because of their reduced coefficient of friction, do not carry their share of the load being applied and thus require the remaining plates to absorb more of the power and heat burden. Thus, there may be an induction period where the preparation of the clutch surfaces later affects clutch failure and this process would place importance on the total time of clutch engagement as previously assumed.

Constant Pressure Drop Across Orifice — The assumption of a constant pressure drop across the orifice during clutch plate closure is obviously only a first approximation. The actual pressure drop is expected to follow a curve similar to that shown in Fig. 6.

If the whole time interval of clutch plate application is important then the assumption of constant pressure drop is permissible. However, if only the clutch plate lockup interval is important then the pressure drop across the orifice can no longer be considered constant but would be expected to fall

[3] Article by H. Kroepelin. Kolloid Zeitschrift, 1929, p. 294.

rapidly from some positive value to zero.

The fact that the calculated failure viscosity value is in reasonable agreement with the determined value is a measure of support for the above assumptions.

Application of Poiseuille-Hagenbach Equation — The assumption that flow through an orifice is identical to flow through a capillary of the same radius is also an obvious approximation since the development of a parabolic flow profile requires a length/radius ratio of 100/1. Although the resultant error in the calculations will be relatively small at viscosities in the order of 5000 cp, it should be noted that the length/radius ratio for the orifice in 1.5/1.

DISCUSSION OF SELBY AND DEEN-STENDAHL PAPERS

V. I. Improver Is Variable In Low-Temperature Viscosity

— C. J. Prizer
Rohm & Hass Co.

THE FLOW PROPERTIES of automatic transmission fluids are near and dear to the hearts of V. I. improver manufacturers since we are assigned a large share of both the credit and responsibility for them. With the development of the Brookfield method, Mr. Selby has placed the measurement of viscosity of automatic transmission fluids at temperatures below 0 F on an objective basis and he has presented convincing arguments for the correlation of viscosity measured in this fashion with the actual performance of a transmission. The composition of a V. I. improver can affect the low-temperature viscosity of the finished fluid. This is a variable to consider carefully in the design of automatic transmission fluids, as illustrated by Table A. Here are shown two typical and fully compounded base oils thickened to 50 SUS with each of two methacrylate copolymers which differ considerably in composition. Notice that with Base Oil A the two fluids have about the same viscosity at – 10 F but at – 40 F. Polymer 2 leads to far better fluidity than Polymer 1. On the other hand, in Base Oil B Polymer 1 gives significantly lower viscosity at – 10 F; and at – 40 F the viscosities are fairly close. In designing automatic transmission fluids, therefore, it is important to keep in mind the variable of V. I. improver composition when low-temperature viscosity is considered.

No Variation Allowable In Brookfield Technique

— A. J. Bozzelli, Jr.
Sun Oil Co.

WE BELIEVE the Brookfield technique is an excellent procedure, but we also feel that each step described has to be followed exactly, if severe difficulty is to be avoided.

Fig. A illustrates this point. These data are the results of Brookfield determinations made on the same fluid, but with *only* a slight variation in procedure. For example, the lower solid line shows the analysis of the fluid according to the procedure described in Mr. Selby's paper. However, when we finished with the analysis at – 40 F we placed the samples back into the cold box, *without re-heating*. This was done at several temperatures with a determination at each temperature. As can be seen, a significantly different curve is established.

We bring this to your attention only as a means of cautioning those who are using the Brookfield technique. We believe it is necessary to follow this procedure without variation, if meaningful results are to be obtained.

Shear Stability Is Prime Quality of Well-Balanced ATF

— C. C. Colyer
Standard Oil Co. (Indiana)

THE PETROLEUM and automotive industries agree that automatic transmissions are one of the most complicated mechanisms in today's passenger cars. Consequently, automatic transmission fluids are among the most complex lubricants used in passenger cars. The authors of these papers have stressed the qualities that an ATF should possess and have presented test methods to evaluate these qualities in order to insure satisfactory field performance. Mr. Selby, by showing that the transmission will fail when a fluid reaches a critical viscosity, has given the petroleum industry an incentive to improve the low-temperature fluidity properties of ATF's. In addition, Mr. Selby has reported good correlation of results from the Brookfield bench test and the Automatic Transmission cold-room test. Thus, the petroleum industry is provided with an available test tool to develop ATF products possessing improved low-temperature fluidity.

Most of our efforts have been in the areas of the overall performance properties required for a well-balanced fluid. Improved shear stability has been of prime interest. I would like to add some confirming data to Messrs. Deen and Stendahl's shear studies with Type A fluids and in

Table A — Effect of Methacrylate Copolymer Composition on Low-Temperature Fluidity of Automatic Transmission Fluid

	Base Oil A		Base Oil B	
	Polymer 1	Polymer 2	Polymer 1	Polymer 2
Viscosity at 210 F, SUS	50	50	50	50
Viscosity at 100 F, SUS	182	184	166	171
Viscosity Index	144	143	152	151
Viscosity at – 10 F, cp	2120	2280	1650	2470
Viscosity at – 40 F, cp	64,000	36,500	42,150	48,750

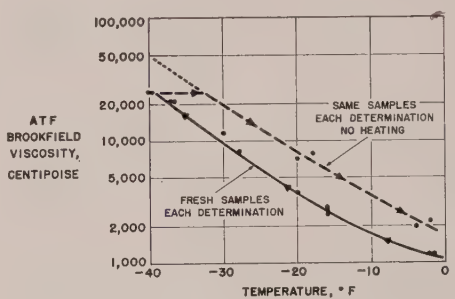


Fig. A — Procedure variation effect

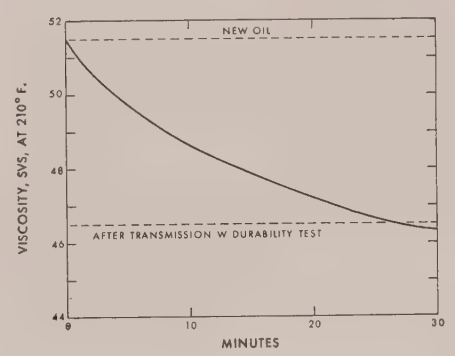


Fig. B — Shearing by sonic oscillation

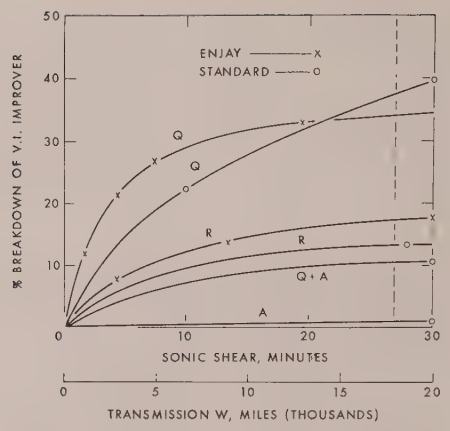


Fig. C — V.I. improver shear stability

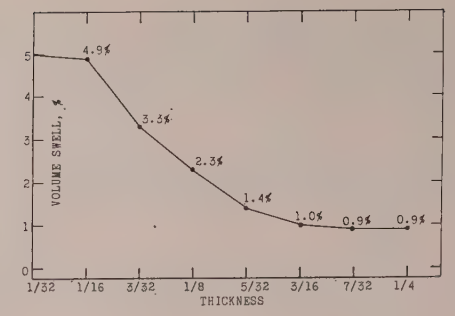


Fig. D — Acadia No. 3612 sheet stock cut ¼ in. wide by 1 in. long cut to various thicknesses (immersed in Type A fluid 70 hr at 300 F)

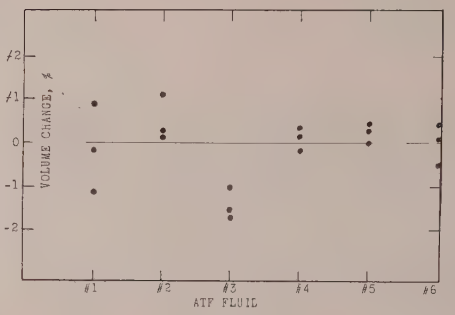


Fig. E — Acadia No. 3612 molded to Chevrolet Part 3702920 seal sections cut to 2¾ in. lengths, immersed for 70 hr at 300 F

addition present a few results from our study of GM's new proposed fluid.

The Raytheon Sonic Oscillator can be used as a fairly reliable shear screening test, with a given laboratory. For example, as shown in Fig. B, a typical Type A-Suffix A fluid of 51.5 SUS at 210 F after 28 min in the Sonic Oscillator shears to 46.5 SUS — the same sheared viscosity the fluid had after being run iin the hydramatic durability test. In Fig. C, Sonic Oscillator results for V.I. Improvers Q and R are compared with results from Transmission W as given in Fig. 13 of the Deen-Stendahl paper. Sonic shear results for a Type A blend containing V.I. Improver Q showed 38% breakdown compared with Enjay's 34% breakdown after 20,000 miles in Transmission W. For a Type A blend containing V.I. Improver B our laboratory obtained 13.5% versus a calculated value of 16% by Enjay. Both laboratories show V.I. Improver R to be much more shear than V.I. improver Q .

The Q and R blends are Type A-Suffix A and neither meets the proposed specification presented by N. A. Hunstad in January, 1959.[a] For the sake of brevity, we commonly refer to this new specification as "Type AA." With V.I. Improver A, which allows little or no shear, the Type AA requirements can be met. A combination blend containing both V.I. improvers A and Q also meets Type AA and allows greater flexibility in selecting base stocks.

Because these V.I. Improvers vary widely in shear stability, the fluids from which they are made will vary widely in initial viscosity. For example, V.I. Improver Q would have to be blended to approximately 68 SUS at 210 F, R to 56.5, A+Q to 55, and A to 52.5 in order to meet the Type AA specification of 52.4 SUS minimum after shear. Automatic transmissions are calibrated at the factory with new fluids. Thus, the more the fluid shears in service, the more the transmission is operating away from optimum calibration and shift efficiency.

Performance Tests Are Final Measure Of Automatic Transmission Fluid

— E. F. Koenig
Texaco, Inc.

A GENERAL COMMENT: Although we recognize the desirability and value of various types of bench tests particularly as research tools in the development of automatic transmission fluids, we should like to extend a word of caution against the overuse of such tests as substitutes for performance tests or in specifications. The overuse of tests of this type in a specification, especially if made too restrictive, can easily eliminate products which would give entirely satisfactory performance in the actual equipment for which they are intended.

Transmission cold-room data which Texaco obtained sometime ago was published in SAE Journal.[b]

The transmission cold-room work which we conducted was somewhat similar to that carried out by Mr. Selby, although it varied in details of operation and interpretation of results. In our work transmission malfunctioning was generally judged when rather severe burning of clutch plates had occurred. In Mr. Selby's work he has tried to look at the plates more critically and judged failure on the slightest indication of incipient burning. In fact his method has sometimes been referred to as "The Sterile" technique. There are no doubt certain advantages in detecting failure at the incipient stage. In this connection we wonder from the data in Table 1 whether the predicted temperature of failure has been selected as being too critical, since all of the predicted temperatures are either equal to or higher than the actual transmission failure temperatures.

We agree with Mr. Selby that low-temperature fluidity is a very important fluid characteristic.

He indicates that there may be other fluid properties and other factors affecting the low-temperature clutch plate failures, but that he considers these definitely secondary to the effect of fluid viscosity and engine torque. For example:

1. As he mentions, it is generally known that certain lubricity and other additives have significant effects on clutch plate coefficient of friction, wear, and general plate life.

2. Some critical moving parts in the transmission undoubtedly have different coefficients of expansion so that the clearance between them may vary with temperature.

We are wondering whether in some instances these factors may not be exerting a greater influence than indicated and would appreciate further comments on these two points.

As Messrs. Deen and Stendahl have indicated, numerous bench tests for measuring oxidation stability of lubricants have been developed and are in use. Each differing in some respect from the others. In fact, there are so many tests of this type available that it sometime appears that it is essentially a matter of personal taste as to which is used. Nevertheless, the test and data presented in this paper are very interesting and probably very useful.

The data on the difference in effect of transmissions on

[a] Letter on "Proposed Automatic Transmission Fluid Specifications," from Norman A. Hunstad, secretary, General Motors Committee on Automatic Transmission Fluid, Jan. 14, 1959.

[b] "Low-Temperature Setup Described," by E. E. Koenig. SAE *Journal*, Vol. 64, November, 1956, pp. 26–27.

the shear down of ATF is also very interesting and confirms opinions and other data that are available. There are also probably considerable data available to show that different transmissions have different effects on other fluid characteristics, such as oxidation stability and deposit formation when operated under similar conditions.

We agree with the authors' comments regarding the repeatability and reproducibility of the ASTM rubber swell tests. As they have demonstrated, correcting for specific gravity of the seal material improves the repeatability of the test. However, there are undoubtedly also other factors of importance in this test. In our laboratory, when we are interested in a direct close comparison, we commonly use the practice of preparing the tests specimens from the same slab of rubber — thus eliminating the variation in specific gravity. Even with this technique it is not uncommon to find variations of from −1 to +1% volume change for a given fluid. At higher volume changes proportionately greater variations have been observed. When different batches of the same type rubber are used variations up to 5% have been observed at the low volume change level (0–7%).

Cooperative Research On Seals Underway

— Bert Vandermar
Acadia Synthetic Products Div.
Western Felt Works

IT IS GRATIFYING to see the continued increase of cooperative effort concerning development and control between suppliers of fluids and seals in the automatic transmission field. This effort is taking the form of joint committees made up of seal and fluid suppliers, as well as automotive manufacturers, direct cooperative evaluations between individual fluid suppliers and seal fabricators, and papers such as this. Not too many years ago, there was little or no liaison between seal and fluid suppliers. The attitude taken was "here is our fluid, adjust your seals accordingly" or, conversely, "here is our seal, adjust your fluid accordingly."

Now we have a committee, Task Group "B" operating within the structure of the technical committee on Automotive Rubber, commonly called "Tech A." Task Group B is made up of representatives of fluid suppliers, seal fabricators, and automotive manufacturers. Our goal is to run exhaustive immersion tests at various temperatures, using representative fluids supplied by the fluid suppliers, through Subcommittee G of the SAE Fuels and Lubricants Committee, and seal stocks supplied by seal fabricators and polymer suppliers. Standard ASTM test procedures are used throughout. All tests are run in triplicate by three separate testing laboratories. Thus far, we have completed a study of gear lubricants and are in the midst of a transmission fluid study.

The data obtained has given indication of the temperature limitations of some seal stocks, as well as some lubricants. Since the study included new high-temperature seal compounds, it gave an insight into future sealing capabilities. This then could be a stimulant for the development of new fluids that were heretofore withheld because present seals were not satisfactory.

Another example of cooperative effort is the direct contact work being carried on. One fluid supplier has no setup for running the "effect on seals" test and so he sends his samples to us for evaluation. Another fluid supplier has established a working agreement with us in which he sends representative samples of any new fluid developed and we, in turn, send him samples of any new seal compound developed. In this way, we are both apprised of new developments in each field.

Open discussions serve to promote understanding of motives and methods. In any compatibility testing, it is obvious that both of the elements cannot change. Therefore, it was necessary some years back to select a standard seal compound for evaluation of any new and different fluids. At that particular time, a considerable number of lip-type seals were being made from Acadia No. 3612; and so it was chosen as a base line stock. There was no intention then, and there is none now, to use the "effect on seals" test as the sole criterion of the quality and performance of a Type A fluid (Fig. D). It was included in the fluid specification as protection for not only the automotive manufacturer, but also for the fluid supplier and seal fabricator, against the use of base stocks of widely divergent aniline points or additives that have a deleterious chemical effect on seals. There have been two specific instances where screening tests of the seal compatibility type have been used to alert automotive manufacturers to potential service problems.

Of course, this test like any other must be used in accordance with standard prescribed test methods — and the results should be interpreted with understanding. Where the test is described, it calls out the use of a specific size sample. Different size samples will give different volume change readings in the same oil at the same temperature because of the ratio of surface area to volume.

Volume swell of a seal stock is dependent on many things: base polymer, plasticizer content, state of cure, and loading. This last, filler loading, affects the specific gravity. High loadings give high specific gravity and correspondingly lower volume swells in a given oil. This point was brought out very well by the authors. As long as rubber compounders have freedom in selection of ingredients and formulation, two or more stocks meeting the same specification will differ in some respects. However, if one compound — in this case No. 3612 — is used in a standard form — in this case Chevrolet part 3702920 — as called out in the ATF specification, quite consistent results can be obtained. These are shown in Fig. E. These results were obtained on six different ATF fluids that are now in production use. The slight shrinkage obtained in some of the fluids can be explained by noting recent trends in aniline points of base stocks. When No. 3612 was chosen some years back, most ATF's were in the 195–198 F range. Now, many of the fluids range as high as 208–210 F. It may be that a new standard seal stock should be chosen. However, this matter is beyond the scope of this discussion.

Finally then, the best course of action is one of cooperative effort between fluid suppliers, automotive manufacturers, and seal fabricators. All three have a common interest in the performance of the transmission.

Mr. Selby's Closure To Discussion

MR. KOENIG asks if, in Table 1, the predicted temperature of failure is too critical, since all the predicted temperatures are either equal to or higher than the actual transmission failure temperatures. The answer is that the predicted temperature of failure must always be equal to or higher than the actual failure temperature if incipient burning is the criterion since the severity of burn on the clutch face is used to determine the predicted temperature of failure. In other words, it would be impossible to judge the temperature of failure from a clear clutch face.

Mr. Koenig also wonders whether the additives in the fluid or the differential expansion of certain critical transmission parts might be exerting a greater influence than indicated in my paper. I can only respond that Fig. 5 shows some variation between fluids which cannot be accounted for by an analysis of known errors. This variation may be due to the causes he has suggested or to others unmentioned; however, at this time it seems from the data collected, that the effects are of a secondary order of magnitude.

I agree completely with Mr. Bozzelli's remarks about the care with which the Brookfield analysis must be obtained. I might add that one must always be conscious of the fact that future fluids and transmissions may require new viscometric tools, techniques, and studies.

High-Temperature and their heat-treatment

O. W. McMullan
Federal-Mogul-Bower Bearings, Inc.

THE HISTORY of the use of metals has been one of ever increasing demand for higher strength and hardness, first for static applications or small tools and later for moving parts in machine components. The application of metals and other materials to elevated temperature use has been no exception to this pattern. Metals were for many years used at high temperatures for such static parts as still tubes, furnace parts, containers, trays, and fixtures before being used in moving parts. The requirements of gas turbine engines have been among the most important factors in bringing about a greater need for materials operating at elevated temperatures. Parts moving at high speeds necessarily require high strength and when in contact with others, a hardness, wear and gall resistance, and dimensional stability beyond the capacities of the more commonly used metals both ferrous and nonferrous. Antifriction bearings are in such a category.

THIS PAPER deals with steel and its heat-treatment employed in antifriction bearings for gas turbine engines. The author classifies the available materials into several groups suitable to various temperature limits ranging from 450 to above 1500 F.

The development and heat-treatment of case-hardened steels are also discussed, along with the metallurgical problems encountered.*

Characteristics and Classification of Steel

When quenched steel is heated the microstructure, hardness, and volume are affected. These three things always occur simultaneously. They are important factors connected with materials and their heat-treatment for performance in bearings. Hardness and strength are, of course, necessary for wear and endurance under moving loads. Volume changes mean lack of dimensional stability. Since clearances in antifriction bearings may be only fractions of one-thousandth of an inch, dimensional stability is important. Heat from whatever source, tempering in the production of parts or in their operation, is the cause for changes in microstructure, and this in turn causes the changes in hardness and volume. Thus, where heat is involved in the application, there may be a closer than normal relationship between materials and their heat-treatment in providing a satisfactory product.

For purposes of this discussion, the available materials for antifriction bearings will be classified into those suitable up to several loosely defined temperature limits as follows:

Up to 450 F
Up to 600 F
Up to 1000 F
Up to 1500 F
Above 1500 F

Applications and Heat-Treatment for Conventional Bearing Steels

For the first group up to about 450 F the conventional bearing steels, such as SAE 52100 and 4620, can be used. At higher temperatures their hardness drops, and they do not have the required

* Paper presented at SAE Annual Meeting, Detroit, Jan. 12, 1960.

Materials

for antifriction applications

wear resistance and endurance properties. The maximum temperature possible depends on loads, speeds and other operating conditions. Since strength varies directly with hardness and, as has been mentioned, hardness and volume changes occur simultaneously, hardness tests can be used as a measure of endurance and of dimensional stability.

Fig. 1 shows in a somewhat exaggerated manner the changes that occur in tempering steel. The specimens were carburized SAE 4620 and direct quenched from 1700 F, hence contain more austenite than normally treated bearing races. They were multiple tempered 1 hr each at temperatures shown with intermittent cooling to room temperature. At all three temperatures, the first tempering causes a drop in hardness by tempering the freshly formed tetragonal martensite. Little or no change occurs by repeating the treatment at 350 F, since this temperature is too low to decompose austenite. At 400 F austenite decomposes slowly and at 450 F more rapidly to bring about an increase in hardness with repetitions up to the point where the added time at 450 F begins to soften the martensite.

The transformation of austenite may be accelerated by lower cooling temperatures in multiple tempering cycles. Fig. 2 shows this effect in cycles of 1-hr heating at 400 F and 1-hr cooling at temperatures from room to –100 F. Relatively mild cooling in ice water has considerable effect, whereas one cycle of cooling to –100 F and tempering at 400 F essentially completes transformation to tempered martensite. The trend of the curves indicates that with sufficient number of cycles all the treatments would eventually produce 100% martensite with a hardness of about 62 R_c.

The lower pair of curves in Fig. 3 show similar changes in hardness over a wider temperature range of tempering. The data are from a 5-in. OD bearing race of SAE 4620 steel carburized at 1700 F, reheated to 1525 F, oil quenched, and ground before tempering. Data are averages from several pieces; one group was given a dry ice treatment and the other was not so treated after the first temper. The significance of these changes is evident in the

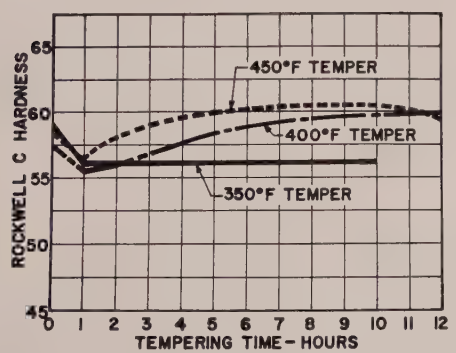


Fig. 1 — Case hardness of SAE 4620 carburized, direct quenched in oil from 1700 F, and multiple tempered as shown. Tempering time shows accumulated hours at heat in cycles of heating for 1 hr each followed by cooling to room temperature

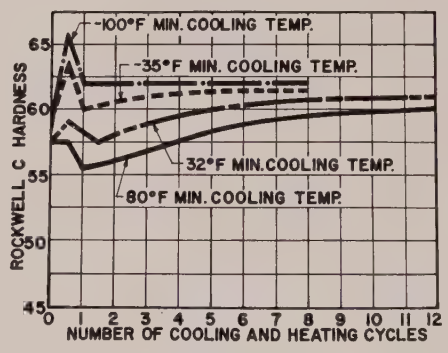


Fig. 2 — Same steel and treatment as in Fig. 1 except for cooling temperatures. All hardness readings taken at room temperature. Hardness after cooling is shown at ½ points. Whole numbers represent both number of cycles and total hours at tempering temperature of 400 F

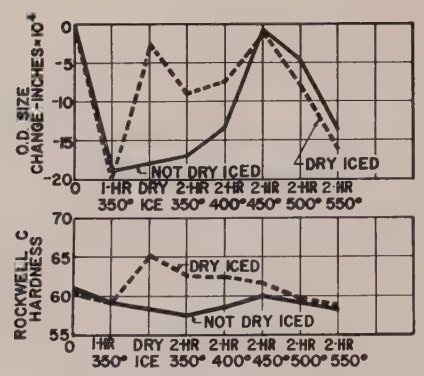


Fig. 3 — Case hardness and dimensional changes of SAE 4620 carburized at 1700 F, box cooled, reheated to 1525 F, and oil quenched in form of bearing race of 5-in. OD. Races ground to size before tempering. Zero point of dimension changes represents size before heat-treatments shown. Dry ice cooled parts to — 90 F

Table 1 — Steels Maintaining R_c of 60 after Tempering at 600 F

Steel	C	Mn	Si	Ni	Cr	Mo	Al
MHT	0.95/1.10	0.25/0.45	0.25/0.55		1.30/1.60		0.75/1.25
TBS-600	0.95/1.10	0.60/0.80	0.85/1.20		1.25/1.65	0.25/0.35	
CBS-600	0.17/0.22	0.50/0.70	0.90/1.25		1.25/1.65	0.90/1.10	
C	0.17	0.55	0.30			2.98	
C-1	0.22	0.57	0.47	1.85		1.94	0.75

upper pair of curves in Fig. 3 which show changes in dimensions that may exceed specified tolerances if normally treated parts are raised to elevated temperatures. Therefore, bearing components must be tempered before final grinding at a temperature above the maximum expected in operation. Since tempering is time sensitive as well as temperature sensitive, either the tempering time or the temperature must be such as to make the parts dimensionally stable for the total time at the maximum operating temperature. Normal tempering time at about 50 F above the maximum temperature during operation has proved to be satisfactory.

Steels for Mildly Elevated Temperatures

For temperatures above those for which such steels as SAE 4620 and 52100 were suitable, two courses were open for bearing producers and both have been followed. One was to select existing compositions that had been used for other purposes at elevated temperatures such as cutting tools, dies, or valves, and the other to develop new compositions for bearing purposes. As a producer of case hardened parts, my company was interested in case hardening compositions, but no such compositions were in common use for other high-temperature purposes; hence, the course of developing new compositions was followed. There has been a general tendency in the bearing industry to select compositions that keep to a minimum the more highly strategic, scarce, and more costly alloying elements. For this reason cobalt and tungsten were not included in some of the investigations described below. As might be expected, the first interest was in mildly increased temperatures representing the second group up to 600 F. This was a natural limit set by characteristics of liquid lubricants.

Preliminary tests indicated the need for more basic knowledge about the effect of individual alloying elements on the resistance of steel to softening by tempering. Fig. 4 shows these effects in case-hardened steels containing increasing amounts of a single-alloying element such as might be required for mildly elevated temperatures represented by the second group up to 600 F. The true effect of nickel after tempering at 350 F is masked by increasing percentages of autenite with increase in nickel content. At 550 F (not shown) and at 700 F it is evident that nickel promotes softening and on a hardness basis is not a suitable alloying element for elevated temperature steels. Other considerations to be discussed later may show favorable effects from nickel. Among those investigated, aluminum, molybdenum, and silicon promoted retention of hardness. Another element not shown on the chart is a very potent one. High carbon content is necessary to obtain and maintain maximum hardness at elevated temperatures. Several low-alloy content steels with carbon contents of 0.70–1.00% were carburized before heat-treatment and found to retain a hardness of 2–4 points Rockwell C higher than uncarburized specimens when both were tempered at 450–550 F. Possibly the most desirable carbon content depends on the kind and amount of alloying elements present, but this has not been determined. Table 1 lists nominal compositions of steels that will maintain a hardness of Rockwell 60C after tempering at 600 F.

Steels for Higher Temperatures

Steels in the third group are those that may be suitable for use in the range of 600–1000 F. Their applications, however, are usually in the range of 400–600 F because of lack of suitable liquid lubricants above 600 F. Most of the compositions presently used are in this group. The chemical analyses of a number of such steels are listed in Table 2.

The suitability of steels in this group for temperatures of 600–1000 F is based on the hardness that will remain after tempering within that range. The significant hardness is the hot hardness or hardness at the operating temperature, but it has been found that there is a fairly close relationship between hot hardness and recovery hardness, that is, the hardness that will remain after returning to room temperature. The necessity for tempering higher than the maximum operating temperature also applies to this group in order to maintain dimensional stability. Hardness is not the only criterion for selecting a steel for use up to 1000 F. For example, all of those listed in Table 2 are subject to oxidation and scale formation unless protected by a lubricant or inert atmosphere.

Stainless Steels

When oxidation may occur or corrosive media are encountered, stainless properties are necessary. SAE 51440 C or modifications are used. The hardness of 440C may drop to 55 Rockwell C or lower when tempered at 1000 F. This can be raised about 4 points Rockwell C by substituting 4% of molybdenum for an equal amount of chromium without losing much in oxidation resistance. Some steel producers have added small quantities of other ele-

Table 2 — Chemical Analyses of Selected Steels

Steel	C	Mn	Si	Ni	Cr	Mo	W	V
T-1	0.65/0.75	0.20/0.40	0.20/0.40		3.75/4.50		17.25/18.75	0.90/1.30
M-1	0.75/0.85	0.20/0.40	0.20/0.40		3.75/4.50	7.75/9.25	1.15/ 1.85	0.90/1.30
M-2	0.78/0.88	0.20/0.40	0.20/0.40		3.75/4.50	4.50/5.50	5.50/ 6.75	1.60/2.20
M-10	0.85	0.25	0.25		4.00	8.00		2.00
M-50	0.80	0.30	0.30		4.00	4.25		1.00
Halmo	0.58	0.30	1.00		5.00	5.25		0.65
VSM	0.70	0.50	1.10		3.00	5.25		
315	0.10/0.15	0.40/0.60	0.20/0.35	2.60/3.00	1.35/1.75	4.80/5.20		

ments, such as vanadium, to the modified composition. These steels are subject to some oxidation as temperatures increase, but have been found useful at 800–1000 F for intermittent or oscillating motions such as in after burners. The optimum amount of molybdenum was reached at 4% in contrast to the 2% illustrated in Fig. 4 where total alloy content was lower. Chemical compositions of 51440C and a modification are given in Table 3.

Table 3 — Chemical Composition of 51440C Steel

Steel	C	Mn	Si	Cr	Mo
51440C	0.95/1.20	1.00 max.	1.00 max.	16.00/18.00	0.75 max.
Modified	0.95/1.20	1.00 max.	1.00 max.	14	4

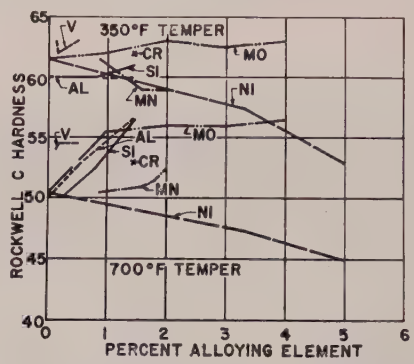


Fig. 4 — Case hardness of steels as affected by different percentages of single-alloying element. Steels carburized 1700 F, box cooled, reheated 1525 F, and quenched. Low manganese steel, used also for zero nickel and molybdenum, was water quenched because of low hardenability. All other oil quenched. All specimens single tempered 2 hr. Al, Si, and V — alloy base; others — carbon base

Nonferrous Metals

Ferrous base alloys become unsatisfactory when temperatures climb above 1000 F. There has been little experience with bearings at such temperatures. The so-called "superalloys" are nonferrous with base compositions of cobalt, chromium, and molybdenum that may also contain nickel, tungsten, vanadium, and others. Numerous compositions exist as proprietary and trade named products and have had some use up to 1500 F or higher. Those of higher carbon content and in hardness ranges normally found in bearings are cast products that cannot be forged or machined. They are cast to close dimensions such as by investment casting. Brittleness increases with carbon content and hardness.

Cermets and Ceramics

Cermets and ceramics constitute the last group. The cermets are cemented carbides. Titanium carbide appears to be the most successful. Tungsten carbide is entirely unsuitable, because it oxidizes and flakes badly when exposed to air at high temperatures. Chromium carbide is rather sensitive to both thermal and physical shocks. Control of the size and shape during sintering are difficult, and diamond grinding is slow and expensive. Ceramics for bearings are in the experimental stage and likewise present problems in production because of brittleness. Aluminum oxide and silicon carbide are materials of this kind. The great differences in thermal coefficients between the cermets and ceramics as compared to shaft and housing materials present problems in design because of the possibility of races cracking on shafts or becoming loose in housings.

Heat-Treatment of High Alloy Content Steels

Mention was made earlier of a desire to use a case-hardened steel for high temperatures. Its development and heat-treatment brought out some metallurgical features of general interest concerning other high-temperature steels for bearings or other purposes. Some of the objectives were to develop a steel that would:

1. Contain only the more readily available alloying elements.
2. Carburize readily to a sufficiently high percentage of carbon.
3. Have a composition that could be carburized and hardened in production-type furnaces, rather than special equipment such as tool hardening furnaces for high-speed steel.

Most of commercially produced high-temperature steels contain various amounts of the carbide-forming elements, chromium, molybdenum, vanadium, and tungsten. They are gamma loop closing elements that raise heat-treating temperatures. Attempts to carburize some of these steels revealed another difficulty. Steels containing as little as 3 or 4% chromium would not carburize uniformly at 1700 F. Some sports had no increase in carbon whatever on freshly cut, clean surfaces, whereas rolled or forged surfaces carburized uniformly. The formation of a green chromium oxide apparently served as an effective stop to carbon penetration.

Because of certain adverse characteristics imparted by aluminum and silicon, molybdenum was chosen as the necessary element to impart hot hardness. Chromium was found to have an additional effect in this combination. With only these carbide formers present, carburizing was slow and

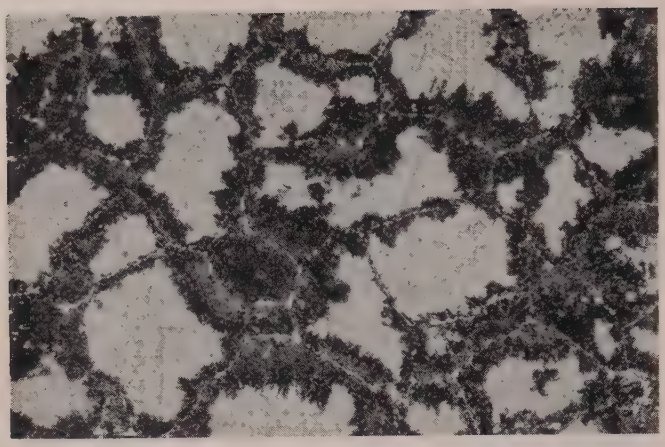


Fig. 5 — Steel 315 with 2.97% Ni, 1.80% Cr, 5.19% Mo. Carburized at 1800 F. Oil quenched and tempered at 1000 F for 100 hr. Nucleation and growth of transformation product starting at grain boundaries in austenitic matrix. Nital etch, original magnified 1000 times

Fig. 6 — Same steel and treatment as in Fig. 5 except multiple tempered four times, ½ hr each at 1000 F with between cooling to room temperature. Light needles are untempered martensite formed below MS point in last cycle. Dark needles are tempered martensite formed in previous cycles. Nital etch, original magnified 1000 times

Fig. 7 — Same steel and treatment as in Fig. 6 except tempered six times to complete transformation of all austenite to martensite. Nital etch, original magnified 1000 times

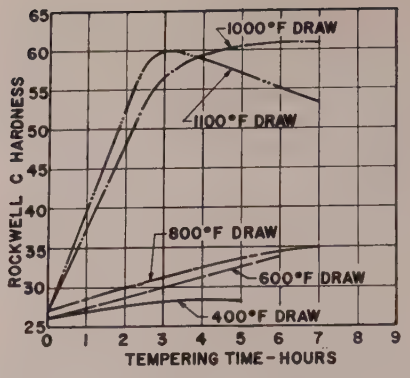


Fig. 8 — Case hardness of Steel 315 carburized at 1800 F, oil quenched, and multiple tempered in cycles of 1 hr at heat, intermittent cooling to room temperature. Tempering time represents both number of cycles and accumulated tempering time

Table 4 — Analysis of Cage Materials

	Cu	Ni	Si	Fe	Zn	Mn
Iron-Silicon Bronze	90.00 min		3.25	1.50	2.75	1.00 max.
H Monel	31	63	3.00	2.00		.75
S Monel	30	63	4.00	2.00		.75

not to a satisfactory carbon content at normal carburizing temperatures. Nickel, the most effective element for widening the loop and decreasing temperatures of austenite formation, was added. Production of a uniform high carbon case was then possible at temperatures of 1700–1800 F from the composition given as 315 in Table 2. This steel has an alloy content sufficient to maintain a hardness of $60R_c$ after tempering at 1000 F. Thus, nickel, not a carbide-forming element and normally considered as an element that decreases the rate of carburizing, was essential in producing a steel of higher alloy content that would carburize under normal conditions.

However, the presence of nickel created another problem. The case tended to be completely austenitic with a Rockwell C hardness of 30 or less even with a carbon content of 1.25–1.50%. The quenched product was in reverse of normal with the case softer than the core as measured by penetration hardness. On the other hand, the surface was such that it could not be scratched by a good file. This appears to be good evidence that, contrary to popular opinion, such an austenitic structure would have good wear resistance. However, it cannot be tolerated in large amounts in bearings for two reasons. It is subject to plastic flow under load resulting in brinelling and dimensional changes and, under heat or stress, will transform to other constituents of larger volume and thus diminish dimensional stability. The elimination of the austenite posed the final problem in the heat-treatment of this case-hardened steel. This problem is one also present even if less evident in many heat-treated steels of high-alloy content. The heat-treatments described below should bring about similar results in other compositions.

Thermal treatments are normally used to decompose austenite. Single tempering, however, re-

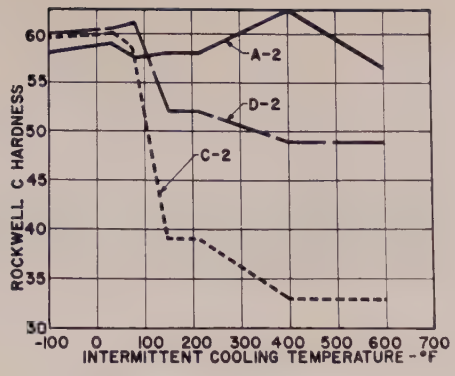


Fig. 9 — Effect of elevated cooling temperatures on start of austenite transformation in multiple tempering cycles. Steels carburized at 1800 F, oil quenched, and multiple tempered at 1000 F with intermittent cooling temperatures as shown. All readings taken at room temperature after completion of last cycle. Steel A2-1.53 Ni, 1.24 Cr, 2.80 Mo; C2-3.06 Ni, 2.85 Cr, 3.12 Mo; D2-2.97 Ni, 1.80 Cr, 5.19 Mo

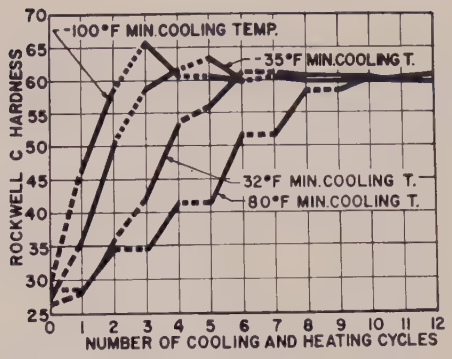


Fig. 10 — Effect of depressed cooling temperatures on rate of austenite transformation in multiple tempering cycles. Steel 315 carburized at 1800 F, oil quenched, and cooling at temperatures shown for 1 hr between heating periods of 1 hr at 1000 F. All hardness readings taken at room temperature. Odd numbers show hardness after cooling, even numbers after heating. Accumulated tempering time in hours is one-half number of cycles

gardless of time or temperature, was unsatisfactory either because of lack of transformation or the nature of the product formed. Periods as long as 100 hr at 1000 F produced only partial transformation by nucleation and growth, a globular decomposition product too low in hardness. Multiple tempering in short periods of not more than 1 or 2 hr each produced an acicular transformation product (martensite) with the required hardness. Fig. 5 represents the microstructure after single tempering at 1000 F for 100 hr; Fig. 6, the partially transformed acicular product after a few cycles; and Fig. 7, the fully transformed product obtained by multiple tempering.

The optimum temperature for multiple tempering depends on the per cent and kind of alloying elements. Transformation curves for Steel 315 are shown in Fig. 8. Temperatures up to 800 F are too low to cause little if any transformation, whereas 1100 F is so high as to cause undesirable softening of the reaction product. The best temperature was found to be 1000 F for this steel.

There has been disagreement as to where transformation takes place — at the tempering temperature or on cooling therefrom. This becomes important in multiple tempering in order to determine the intermittent cooling temperature between heating cycles. Because of the high percentage of austenite in the quenched 315 steel, the transformation could be readily followed. Fig. 9 shows that the conditioning effect in multiple tempering occurs only after cooling to some minimum temperature between heating cycles. This minimum depends on the stability of the austenite. Steels with stable austenite such as in high-speed and other steels should be cooled to room temperature between heating cycles.

Another observation disputes the generally held impression that holding for a period of time or tempering a quenched steel prior to deep freezing stabilizes the austenite. That appears to be only a special case, whereas the opposite or promotion of transformation is the general situation. If parts containing austenite are deep frozen repeatedly with no intervening tempering, little or no transformation takes place after the first cooling. When tempered between freezings, both the heating and cooling causes further transformation which then proceeds to completion in a relatively short time, as shown in Fig. 10. It is believed that this can be explained by a stress relief theory and that stabilization and conditioning are merely different manifestations of the same thing.

Retainer Materials

Since sliding friction occurs between rollers or balls and cage pockets and between cages and ribs of races, it is desirable to use cage materials that do not wear readily or seize in sliding contact with hardened steel.

A copper-base alloy, most frequently referred to as an iron silicon bronze, has the widest usage for temperatures up to 600 F, or about the maximum for liquid lubricants. This includes the majority of high-temperature bearings. The machinability and bearing properties in sliding friction against steel are reasonably good. It has been used in the cast condition, but more frequently as forged. For higher temperatures, the nickel base monels, H and S, are favored. Both are cast alloys and are used as centrifugally or possibly precision cast. H monel has higher ductility and is used when certain cold forming operations are required. Table 4 lists the major constituents of the three most commonly used cage materials.

Stainless-steel and irons have received some consideration. Nickel-base alloys or some of the more complex alloys containing cobalt, molybdenum, chromium, and other refractory metals offer possible choices for the highest temperatures. Cage pockets are held to close dimensions and many of the alloys present difficulties in machining such as in broaching the pockets.

Several means have been used to reduce friction losses between retainers and the races, balls, or rollers. The most common in connection with liquid lubricants has been silver plating or silver and lead plating. Other coatings, impregnated or plated, have employed graphite, chromium, molybdenum sulfide and oxide films. These may serve as protection against oxidation as well as lubricants at temperatures above 600 F.

TURBINE FUEL

CFR Coker

J. D. Rogers, Jr.
E. I. du Pont de Nemours & Co., Inc.

FUEL THERMAL STABILITY at the operating temperatures encountered in commercial transports and military weapon systems is an important safety and performance consideration, and yet, undoubtedly, one of the most difficult jet fuel properties to measure. Thermal stability measurements are related to the amount of deposits formed on heating fuel in a jet aircraft and in the engine prior to and during the time it is burned.

New research techniques have been employed to determine the threshold temperatures at which the deposit-forming reactions may occur in fuels. These techniques have been applied extensively in laboratory and flight programs to establish standard procedures for use by the aviation and petroleum industries. The CFR Fuel Coker developed by the Coordinating Research Council is the best-known laboratory method relating fuel stability to service experience.

Current jet fuels are thermally stable at 300/400 F. Fuels that will be thermally stable at higher temperatures will be needed for jet-powered aircraft flying at Mach 3 and above, and evaluations of hydrocarbon fuels are being extended into such areas of superstability.

This paper is a report of the Fuel Thermal Stability Group, Aviation Fuel, Lubricant, and Equipment Research Committee, Coordinating Research Council, Inc.*

THE PREMATURE deterioration of liquid hydrocarbons presents a problem of widespread concern, and petroleum fuel research has been directed over the years toward a better understanding of the decomposition of gasoline, distillates, and jet fuels. Traditional laboratory methods have not been very reliable for predicting the stability of jet fuels. No correlation was found between filterability, peroxide number, and insoluble residue formed in storage,[1] or by accelerated oxidation bomb tests. Evaluating the effects of elevated temperatures on jet fuels is a formidable task. Bromide number, olefins, aromatics, sulfur, and gum values of jet fuels offer no real clues to their stability. New techniques had to be established in order to cope with problems at high temperatures. In 1955, the Coordinating Research Council was assigned the responsibility of developing valid laboratory procedures, and this effort bore early fruits.

A metal tube and filter unit soon became the hallmark of high-temperature fuel research. A significant breakthrough was accomplished by standardizing this equipment in the CFR Fuel Coker.[2]

This report presents the results of investigations leading to (1) research methods for determining the stability threshold temperature of jet fuels; (2) correlations of standard CFR Fuel Coker results, flight tests, and service experience; and (3) research techniques to explore the stability of fuels at bulk temperatures above 500 F.

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 7, 1959.

[1] "Jet Fuel Storage Stability." CRC Project No. CG-1-58, April, 1958.

[2] "Tentative Method of Test for Thermal Stability of Aviation Turbine Fuels." ASTM designation D1660-59T. ASTM Annual Meeting, June, 1959.

THERMAL STABILITY

and Flight Evaluations

Valuable information has been developed with new fuel-sampling techniques and research methods[3] for measuring the stability threshold temperature. Flight test data and an analysis of engine operation under service conditions established a direct correlation between the laboratory fuel ratings and full-scale engine performance.[4] The design of a research coker with a heated fuel reservoir and increased heating capacity in the preheater section provides high temperatures for brief periods to permit evaluations at conditions encompassing most of the advanced propulsion systems.[5]

Background

Relatively low volatility is desired to minimize jet fuel evaporation and precipitation of insoluble residues. In addition, the capacity of the fuel as a coolant is important. In pressurized liquid fuel systems, smaller heat transfer surfaces and lower circulating rates can be used with fuels of high specific heat. Furthermore, thermally stable fuels mean less weight for insulation or refrigerating equipment, and thus more range or payload.

In supersonic flight, fuel is often used to cool the engine oil by heat exchange. However, if fuel deposits accumulate in the heat exchanger and in the burner nozzles, severe engine damage can be encountered.[6] The time factor is important to considerations of flight safety. Fuel nozzle deposits may occur gradually and then increase rapidly, causing uneven burner temperature patterns which distort turbine components to the point of failure. The performance of the engine, as well as the life of critical components required to keep the engine in prime operating condition, is fundamentally related to the thermal stability of the fuel.

The occurrence of deposit-forming reactions may have a pronounced effect in reducing the residence time at which fuels can be held at elevated temperature, as illustrated in Fig. 1. The slope of the temperature-time relationship is a complex function which may vary significantly with different fuels. As an example, Fuel A is less stable than Fuel B in a propulsion system, even though the residence time is 45 sec or less from the time the engine pump picks up the fuel until the fuel is burned. Furthermore, reactions initiated during weeks or months of ambient storage and accelerated by temperatures approaching 285 F for several hours in the airplane tank may form heavy deposits in the engine fuel system.

Typical carbonaceous deposits removed from jet engine fuel nozzles have been analyzed for trace elements.[7] Metals are usually present, and the organic part of these deposits contains a large proportion of sulfur, nitrogen, and oxygen. Oxygen from air which is generally absorbed during fuel processing, storage, and testing promotes the for-

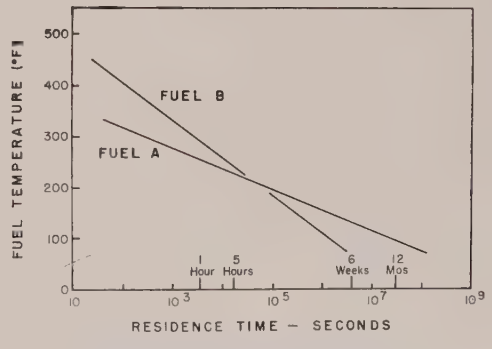


Fig. 1 — Temperature-time relationship for fuel stability

[3] "Fuel Thermal Stability Exchange Program." CRC Project No. CA-2-58, September, 1958.
[4] "Investigation of Thermal Stability of Aviation Turbine Fuels with CFR Fuel Coker." CRC Project No. CFA-2-54, April 2. 1957.
[5] "High Performance Hydrocarbon Fuels for Supersonic Propulsion," by M. E. Conn and W. G. Dukek, Jr. Paper presented at SAE National Aeronautic Meeting, New York, April, 1959.
[6] "Relation of Fuels to Aircraft Turbine Engines," by E. A. Droegemueller and R. K. Nelson. Paper presented at SAE National Aeronautic Meeting, New York, April, 1958
[7] "High-Temperature Stability of Jet Fuels," by G. F. Jonke and R. K. Nelson. United Aircraft Corp. Report FL-55-13, Nov. 22, 1955

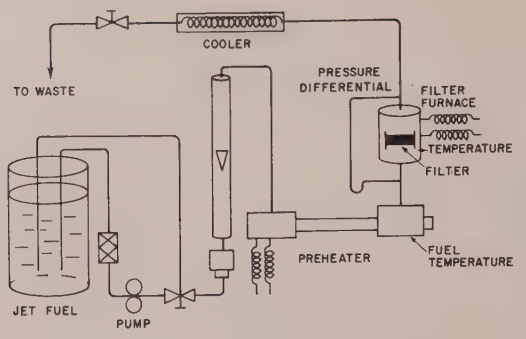


Fig. 2 — CFR Fuel Coker flow diagram

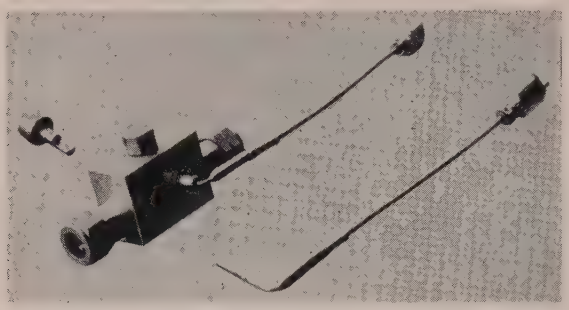


Fig. 3 — Locating gage, preheater, thermocouple

Table 1 — Present and Future Requirements for Hydrocarbon Jet Fuel Thermal Stability

	MIL-J-5624		MIL-J-25656	Future
	JP-4	JP-5	JP-6	JP-X
Test Type	CFR Fuel Coker		CFR Fuel Coker	CFR Research Coker
Test Temperatures, F				
Preheater	300	300	400	500
Filter	400	400	500	600
Time, min	300	300	300	300
Fuel Flow Rate, lb/hr	6	6	6	—
Maximum Filter ΔP, in. of Hg	12	12	10	1
Maximum Preheater Deposit	3	3	2	1

mation of deposits,[8,9,10] particularly in the presence of copper and certain reactive sulfur and nitrogen compounds.

Despite the limitations of routine aging tests [11,12] the deposits from jet fuel appear to be related to gasoline gums and the insoluble residues precipitated from fuel oils upon long periods of storage.

The technical background contributed by the petroleum industry has enabled CRC to set forth realistic working principles to guide the current and future investigations. The stability targets for present and future fuels as defined by the Air Force are shown in Table 1.

Standard CFR Fuel Coker

The CFR Fuel Coker technique for evaluating the thermal stability of aviation turbine fuels was accepted by the ASTM as a tentative standard at the annual meeting in June, 1959, and is now under the jurisdiction of Committee D-2 on Petroleum Products. The ASTM Method D1660–59T had been published previously for information only, and the latest modifications recommended by CRC have been incorporated in the new tentative standard.

The equipment for measuring high-temperature stability is shown schematically in Fig. 2.

Fuel is pumped in the liquid phase through an electrically heated preheater secion. Laminar flow is maintained at low Reynolds number in the annular space between the inner and outer tube. The fuel then passes into a heated filter section where fuel degradation products may become trapped.

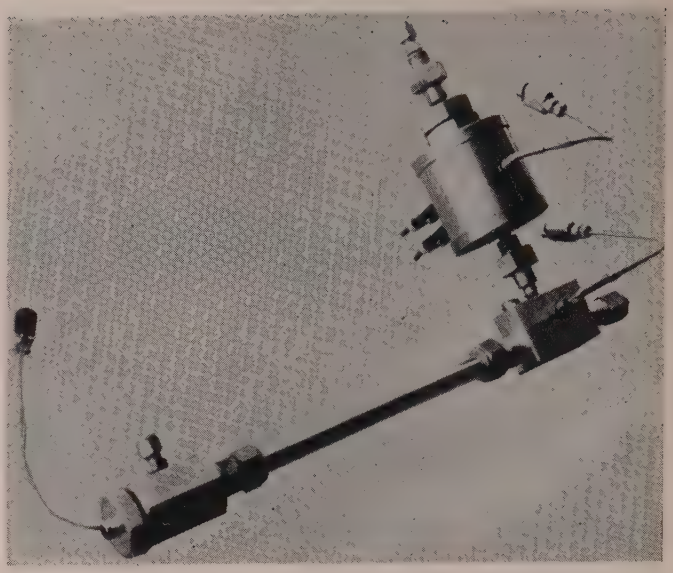
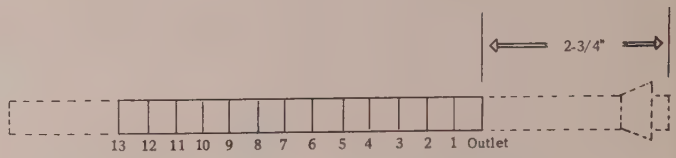


Fig. 4 — Cast-in filter furnace and preheater assembly

Table 2 — CFR Fuel Coker Preheater Tube Deposits



Highest Number in any of above 13 squares
Deposit Rating Code: Use with CRC-ASTM color standards
0 — No visible deposits
1 — Visible haze or dulling, but no visible color
2 — Barely visible discoloration
3 — Light tan
4 — Heavier than Code 3

[8] "Effect of Composition and Storage on Laboratory Properties of Jet Fuels," by A. C. Nixon, C. A. Cole, and H. B. Minor. Paper presented at SAE Summer Meeting, Atlantic City, June, 1955.
[9] "Novosti Neftyanoy Tekniki," by Ya. B. Zherkov, V. M. Shchagin, and V. V. Rudakov. *Neftepererabotka,* No. 2, 3, TsNIITYeneft[1], 1956.
[10] Article by E. G. Davidson. Schweitzer Archiv für Angewandte Wissenschaft U. Technik, Vol. 20, No. 4, 1954, p. 123.
[11] "Studies on Storage Stability of Distillate Fuels," by F. G. Schwartz, C. C. Ward, and H. M. Smith. SAE *Journal,* Vol. 62, February, 1954, pp. 107–109.
[12] "Accelerated Aging Test Limitations," by C. A. Duval. *Petroleum Refiner,* Vol. 30, February, 1951, pp. 85–86.

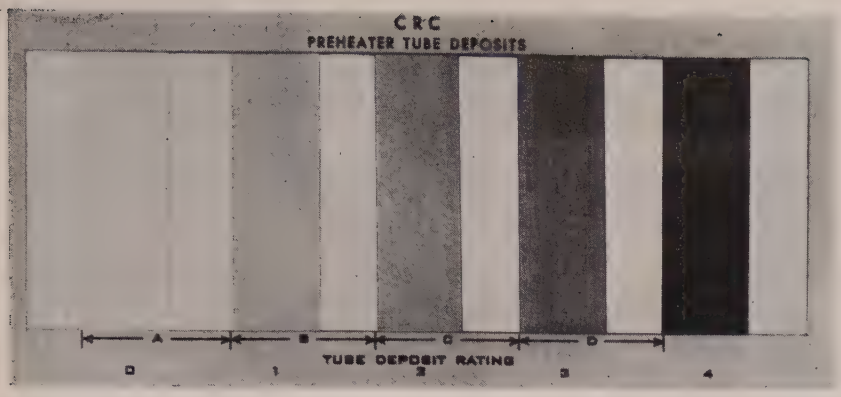


Fig. 5 — Color standards

The stainless steel filter is a sintered matrix with an average pore size of 20 microns. The fuel-flow rate is held constant, and the increased pressure differential (and $\sqrt{\Delta P}$) at the test filter is a direct indication of the amount of deposits being collected. This measurement is used in combination with the deposit condition on the surface of the aluminum inner tube in assessing the thermal stability of the fuel.

The positioning of the thermocouple element in the outlet end body of the preheater unit is essential to making correct readings of fuel temperature. This small (1/16-in. diameter) thermocouple is placed in the centerline of the preheater outlet using the gage shown in Fig. 3.

The filter unit has a cast-in heater and a thermocouple embedded in the metal wall. The assembly of the preheater and filter furnace is shown in Fig. 4.

Rating Preheater Deposits — The preheater deposits are rated visually, after the tube is dried with a stream of pentane or hexane. The tube surface area is examined under a controlled light source at each inch along the tube. The colors of the deposits are matched against the standards shown in Fig. 5 and reported by the code shown in Table 2. The darkest deposits are reported, rather than the average. Occasionally, small individual spots, streaks, or bands of dark deposits are formed, possibly as insoluble residues precipitated by evaporation or nucleate boiling of volatile fuel components at the tube surface. Small individual spots less than 0.05 sq in. should be ignored, provided there are no more than three on the entire tube.

Sampling Techniques

Thermal stability is influenced by the chemical structure of the fuel and by the presence of soluble and insoluble contaminants. The CFR Fuel Coker is very sensitive to fuel contamination and other factors related to deposit-forming reactions, and the effects of storage and fuel contamination become more severe as temperature increases. Consequently, it is important to use clean fuel for the thermal stability measurement. To accomplish this, the fuel is examined for cleanliness and freedom of moisture, and any particulate matter larger than 4 microns is removed by paper filtration. By working with materials of good purity, the thermal stability measurement becomes essentially a function of the chemical structure of the fuel.

The amount of oxygen dissolved in the fuel can also exert a pronounced effect on stability. Therefore, all fuel is saturated at room temperature with air free of oil and moisture. If the fuel has been stored under an inert gas blanket, the aeration time is normally extended from 3 to 20 min.

Theshold Temperature Techniques

The first experiments with a prototype fuel coker were conducted by the CRC Group's Program and Analysis Panel. The Panel employed a 3×3 symmetric orthogonal square and developed statistical information on nine test fuels in 27 cooperating laboratories.

As a result of this phase, modifications in the apparatus and method were evolved. The modified unit was then used to determine reproducibility. The modifications included placing the test filter in a horizontal position and a warmup schedule to control the initial heat input.

A second series of experiments was then conducted with the same nine test fuels and an additional one from the Air Force. These 10 fuels represented commonly available types selected by the Fuels Panel from East Coast, Mid-Continent, Los Angeles, Middle East, and San Joaquin crude sources. The fuels were prepared by conventional refining processes[13,14,15,16] such as acid or solvent treating, used in combination with additives[17,18] to provide a range of thermal stabilities.

Military MIL–J–5624 fuel specifications were followed as closely as possible in defining the JP-4, JP-4 referee, and JP-5 types of test fuels, and a specifically treated kerosene of good stability was also included. The majority of the fuels contained conventional antioxidant and corrosion inhibitors. The gum values were determined on the fuels

[13] "Thermal Stability — New Frontier for Jet Fuels," by A. B. Cramption, W. W. Gleason, and E. R. Wieland. Paper presented at SAE Summer Meeting, Atlantic City, June, 1955.
[14] "Stability to Burn," by D. P. Heath, C. W. Hoffman, and J. H. Reynolds. Paper presented at SAE Summer Meeting, Atlantic City, June, 1955.
[15] "Stability of Aircraft Turbine Fuels," by C. R. Johnson, D. F. Fink, and A. C. Nixon. *Industrial and Engineering Chemistry*, Vol. 46, October, 1954, pp. 2166–2173.
[16] "Oxidation Stability of Aircraft Turbine Fuels," by J. W. Thompson. *Oil and Gas Journal*, Vol. 53, Nos. 5, 7–10, June 7–July 12, 1954.
[17] "Improving Fuel Oils Through Use of Additives," by W. deB. Bertolette and J. D. Rogers. *Petroleum Engineer*, Vol. 28, November, 1956, pp. C41-42, 45-46.
[18] "High-Temperature Stability of Jet Fuels," by C. M. Barringer, M. W. Corzilius, and J. D. Rogers. Petroleum Processing, Vol. 10, December, 1955, p. 1909–1911.

Table 3 — CRC Threshold Stability Test Fuels, Chemical and Physical Properties

CRC-RAF Number	98-55	99-55	100-55	102-55	103-55	104-55	105-55	106-55	107-55	108-55
Fuel Type	JP-4 Ref.	Treated Kerosene	JP-5	JP-4	JP-4 Ref.	JP-5	JP-4	JP-4	JP-4	JP-4 Ref.
Crude Source	—	—	East Coast	Mid-Continent	—	Los Angeles	Middle East	Mid-Continent	San Joaquin	—
Additive Concentration, weight %	—	—	—	0.03 Exp. J	0.01 Exp. K	—	—	—	—	0.04 Exp. L
Antioxidant	0.003	—	0.003	—	—	—	0.003	—	0.003	—
Corrosion Inhibitor	0.012	—	0.01	—	—	0.008	0.01	—	0.008	—
Gravity, deg API	48.5	48.1	44.6	54.1	48.5	37.6	50.5	53.9	47.3	52.5
Mercaptan Sulfur, weight %	0.001	Nil	0.0002	0.00	0.001	0.001	0.001	0.00	0.0002	0.00
Total Sulfur, weight %	0.193	0.016	0.067	0.107	0.195	0.33	0.16	0.08	0.06	0.05
Smoke Point, mm	22.5	34	26.1	29	23	20	27.7	27	22	37
Freezing Point, F	L-67	—	-58	L-67	L-67	L-40	L-76	L-67	L-76	—
Aniline-Gravity Constant	6474	8571	6717	7140	6421	5038	—	7115	5568	6799
Aromatics, volume %	14.6	2.5	10.3	14.5	15.1	19.9	14.4	13.9	12.5	9.1
Olefins, volume %	1.2	5.6	2.7	0.08	0.9	0.8	1.6	1.0	1.2	1.4
Bromine Number	—	—	2.5	—	—	3.2	1.8	—	0.7	—
Existent Gum, mg/100 ml										
Initial	1.4	0.6	0.5	4.0	23.2	0.6	0.4	3.0	1.0	—
Average	1.7	0.7	0.6	6.0	30.8	0.9	0.8	0.8	1.1	13.6
Potential Gum, mg/100 ml										
Initial	9.6	1.2	0.8	9.4	16.0	0.6	0.5	13.2	1.0	25.0
Average	5.1	1.1	0.9	8.0	27.1	2.2	1.6	5.2	2.1	15.6

Table 4 — CRC Flight Test Fuels, Physical and Chemical Properties

	JP-4 Fuels				
CRC-RAF Number	112-57	117-57	135-58	134-58	135-58[a]
Flight Test Number	1	2	3	4	5
Distillation					
Ibp, F	140	148	142	156	—
Evaporation at 290 F, %	51	50	70	78	34
Evaporation at 370 F, %	94	79	96.5	97	66
Evaporation at 400 F, %	98	86	98	99	77
Evaporation at 470 F, %	99	97	99	99	100
End Point	416	492	410	384	470
Gravity, deg API	53.5	52.3	54.9	51.2	50.2
Existent Gum, mg/100 ml	0	4.8	1.0	2.8	1.4
Potential Gum, mg/100 ml	5.0	12.8	14.4	16.5	4.7
Total Sulfur, weight %	0.126	0.052	0.022	0.042	0.1947
Mercaptan Sulfur, weight %	0.002	0.002	0.001	0.001	—
Reid Vapor Pressure, psi	2.6	2.9	2.9	2.8	2.6
Freezing Point, F	- 80	- 90	- 80	- 77	- 76
Net Heating Value, Btu/lb	18,689	18,604	18,622	18,412	—
Aromatics, volume %	9.0	14.9	16.9	24.3	16.4
Olefins, volume %	2.4	4.6	1.1	4.4	0.09
Smoke Point, mm	29	29	33	24	26.8
Smoke Volatility Index	70.2	65.9	74.6	65.6	60.8
Corrosion	0	1a	1a	1a	Pass
Water Reaction	1	1	1	1	1b
Corrosion Inhibitor, lb/1000 bbl	5	5	—	—	4.5
Antioxidant, lb/1000 bbl	8	8	8	6	—
Components					
Virgin Naphtha, %	100	79	81	71	—
Catalistic Cracked Naphtha, %	0	21	0	0	—
Thermal Cracked Naphtha, %	0	0	19	24	—
Commercial Toluene, %	0	0	0	5	—

[a] Strategic Air Command fuel from service.

initially and also later during the test program. The physical and chemical fuel properties pertinent to stability considerations are summarized in Table 3.

A range of temperatures was employed to determine the temperature at which either filter plugging or preheater deposits occurred. Each test was conducted for 5 hr at preheater temperatures varying in increments of 25 F from 275 to 450 F. In most cases, the filter temperature was 100 F higher than the preheater temperature. The filter thermocouple measures a metal-wall temperature, while the preheater thermocouple measures a fluid temperature. This control of the filter gives an indication of temperature approximating that of the liquid fuel at the preheater outlet.

The threshold temperatures determined for the test fuels are shown in Fig. 6. The majority of these test fuels appear to be limited in the 300–450 F range by tube deposits rather than filterability. The average standard deviation of the results was between 9 and 53 F in a few instances, but was between 20 and 37 F for several of the test fuels. Therefore, ± 31 F appears to be a reasonable value for the precision of the threshold temperature determinations.

Fuel instability was indicated arbitrarily by a pass-fail point of No. 3 or higher for preheater deposits and a pressure differential of 13 in. of Hg or higher for filterability. Instability increased as the temperature increased until a threshold was reached. Some of the fuels broke sharply, and the shape of the individual curves resembles an oxidation induction period. Consequently, for a fixed time and temperature, the test is inherently a pass-fail criterion. Stable fuels can be distinguished with a high degree of confidence, and in this regard the test is most effective as a fail-safe device for screening fuels.

A comparison of the threshold stability temperatures with fuel inspection data indicates little, if any, correlation with bromine number, olefin, aromatic, or sulfur content and gum value. Potentially

Table 5 — Thermal Stability of Flight Test Fuels
CFR Fuel Coker: 300/400 F /6 lb per hr

CRC-RAF Fuel Number	112-57	117-57	133-58	134-58	135-58
Flight Test Number	1	2	3	4	5
Time, min	300	300	60	300	95
Filter ΔP, in. of Hg[a]	0.8 (6.5)	11 (24.4)	25 (21.4)	0.25 (3.9)	25 (25)
$\sqrt{\Delta P}$	0.89 (2.5)	3.31 (4.9)	5 (4.61)	0.5 (1.97)	5 (5)
Preheater Deposit	1 (1.4)	3 (2.5)	1[b] (2)	3[c] (4)	4 (4)
Deposit Factor	1.89 (3.9)	6.31 (7.4)	6 (6.61)	3.5 (5.97)	9 (9)
Threshold Temperature, F[a]					
Filter	300	300	225	300	275
Preheater	325	275	275	275	225

[a] Average data obtained periodically during each flight test by WADC.
[b] Code 4 preheater deposit in 300-min test on filter bypass.
[c] Code 4 by CRC-ASTM color standards.
Note: Average CRC exchange data often secured some time after flight test are designated by ( ).

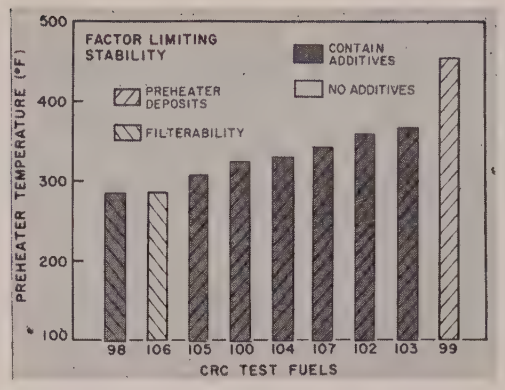


Fig. 6 — CFR Fuel Coker stability threshold temperature

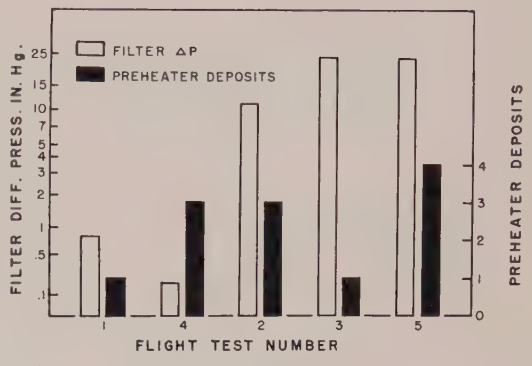


Fig. 7 — Flight test fuels, average CFR Fuel Coker ratings

unstable fuels, for example, may be high or low in gum value. The existent gum values for fuels 98, 99, 100, 102, 104, 106, and 107 were no more than 6 mg/100 ml. The reproducibility of a gum determination in this range is such that any value may be in error by 100%. It must also be borne in mind that the test fuels were stored in different areas and, hence, aged under different climatic conditions prior to the CFR Fuel Coker tests. This introduced an uncontrolled factor, the importance of which cannot be determined from the gum values.

Reference Fuels — Early in 1957, the Air Force requested a reference fuel to be used in calibrating the CFR Fuel Coker. A fuel like RAF–99–55 in the threshold temperature program might be suitable for this purpose, but it was felt that more information is needed regarding its performance after long-time storage, and further work is under way by the ASTM Technical Committee J on Aviation Fuels.

Ideally, two reference fuels essentially chemically pure and stable in storage are desired for rating test fuels, provided unstable and stable ones can be found so that ratings can be expressed in terms of a percentage blend of the two. Work by the Fuels Panel is continuing in CRC to develop information on a stable fuel treated with imidazoline as a means of varying the stability.

In the interim, a program for periodic exchange of fuel samples was instituted under an Exchange Panel to provide equipment standardization data on a variety of fuels. Each month, a sample test fuel supplied by one of the laboratories has been evaluated by the other cooperating laboratories. This work has continued for 21 months, involving the exchange of samples among 40 laboratories. The agreement among participating laboratories in rating stable and unstable fuels at temperatures used by the Military for JP-4, JP-5, and JP-6 fuels has been satisfactory.

Military Service Experience

Air Force Flight Tests — In order to establish correlations between full-scale engine results and CFR Fuel Coker measurements, flight tests were initiated with several test fuels. The Air Force conducted 100-hr flight tests on five fuels having different thermal stabilities as measured by the CFR Fuel Coker.

A summary of the test fuel properties is presented in Table 4, and the CFR Fuel Coker results are shown in Table 5. The stability ratings are average values obtained by WADC at periodic intervals during each flight test. CRC data were also obtained by the Exchange Panel; however, this information was frequently obtained some time after the flight test. The CFR Coker results shown in Fig. 7 represent the average values obtained by WADC.

An F-100 fighter aircraft equipped with a J–57–21 model engine was used for the flight tests. The average temperatures during each test at the oil-cooler outlet before fuel enters the engine manifold are summarized in Table 6, as well as the fuel system cleanliness after each test as determined by nozzle deposits and flow calibrations.

There was no visible distress in the engine fuel manifold with the first flight test fuel, which was stable at 300/400 F in the CFR Fuel Coker. The fuel manifold tubing was clean, and there was only a slight decrease in fuel flow.

Flight test fuel No. 2 was unstable as measured by both filterability and preheater deposits in the CFR Coker. Nineteen of the 48 engine fuel nozzles had sufficient deposits in the main fuel passages to re-

Table 6 — Fuel Performance Evaluations after 100-Hr Flight Tests
(F-100C Aircraft)

Flight Test Number	1	2	3	4	5
J-57 Engine Conditions					
Temperatures, F					
Fuel at Oil Cooler Outlet	139	142	112	132	132
Turbine Exhaust (maximum)	1150	1205	1090	1140	1210
Turbine Exhaust, ΔT	85	95	70	45	115
Fuel System Cleanliness					
Manifold Tubing	Clean	Light Coke	Light Varnish	Moderate Varnish	Heavy Varnish
Number of Nozzles with Visible Deposits	0	19	21	9	35
Plugged Main Screens	0	16	0	0	34
Plugged Swirl Slots	0	4	17	5	21
Poor Spray Pattern	0	5	8	6	13
Fuel Flow Decrease, %	2.44	7.75	8.15	4.70	18.44

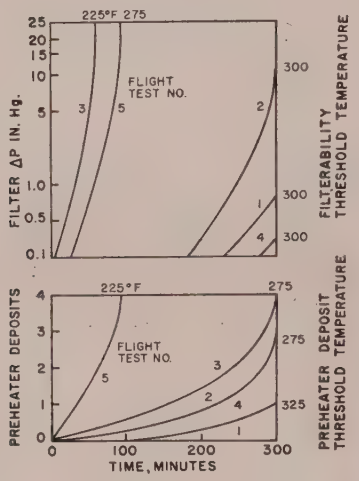


Fig. 8 — Flight test fuels, CFR Fuel Coker ratings

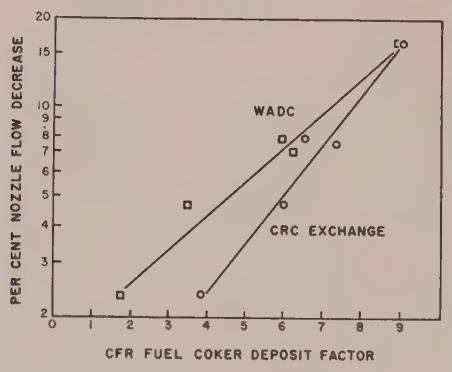


Fig. 9 — Coker results related to flight tests

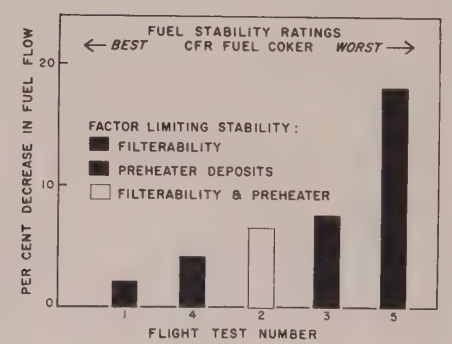


Fig. 10 — Decreased engine fuel flow after 100-hr flight test

duce the overall fuel flow 7.75%. Light coke deposits accumulated in the manifold tubing.

Flight test fuel No. 3 was unstable as indicated by filterability in the CFR Fuel Coker, Twenty-one engine fuel nozzles were plugged in the main fuel passages, and there was an 8.15% reduction in total fuel flow. Light tan deposits in the manifold tubing were similar to a No. 3 CRC color code.

Flight test fuel No. 4 formed moderate preheater deposits in the CFR Fuel Coker. The engine fuel nozzle condition was inferior to flight test No. 1. There was moderate varnish in the fuel manifold tubing and a 4.7% reduction in fuel flow.

Flight test fuel No. 5 was a JP-4 fuel obtained from a western Air Force base which had reported serious difficulties with engine fuel system deposits. With this fuel, 35 engine fuel nozzles were plugged in the main fuel passages; 34 main fuel screens were plugged, and the fuel flow was reduced 18.44%. In the CFR Fuel Coker, it formed both filter and preheater deposits after 95 min.

The threshold temperatures for these fuels indicate that preheater deposits and filter deposits are both important in terms of engine performance. The threshold temperatures determined by WADC are shown on the top and side of the charts in Fig. 8. At or near the threshold temperature, the filter

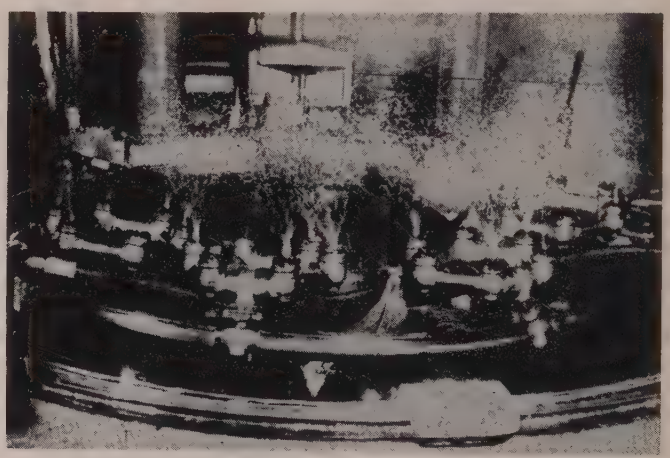


Fig. 11 — Poor fuel nozzle patterns on manifold calibration

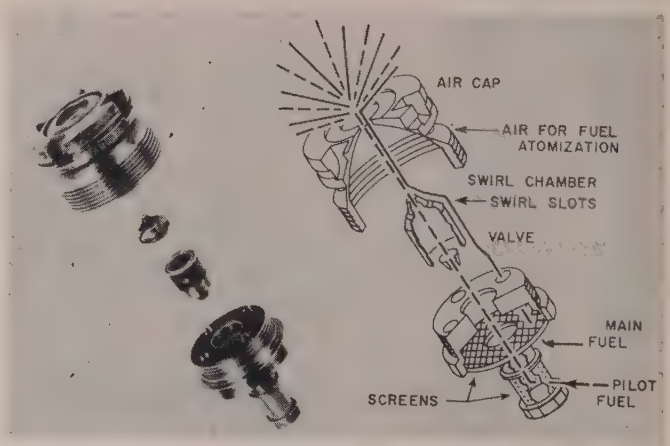


Fig. 12 — Burner fuel nozzle flow diagram

ΔP may become too high to measure accurately. Greater precision may be obtained for research purposes if the arbitrary pass-fail point used for threshold determinations is selected at a lower value of ΔP.

A distinct relation between the flight test data expressed as the per cent decrease in engine fuel flow and the CFR Fuel Coker is obtained by combining the preheater tube deposit rating and the filter rating:

$$\text{Deposit Factor} = \text{Preheater Deposit Rating} + \sqrt{\text{Filter } \Delta P}$$

This relationship is illustrated in Fig. 9. The WADC data show up somewhat milder on the average than CRC Exchange data, which were often obtained some time after the flight tests. The per cent decrease in fuel flow obtained with each of these fuels shows a good correlation with their stability in the CFR Fuel Coker, as presented in Fig. 10.

Air Force Service Experience

The JP-4 fuel used in the fifth flight was unstable at 300/400 F in the CFR Fuel Coker and caused serious difficulties in B-52 jet aircraft operated by the Strategic Air Command. Heavy deposits accumulated in J-57 engine fuel systems in less than 100 hr of service. Premature engine overhauls were necessary in 37% of the engines because of excessive temperature differentials across the turbine exhaust section. This rejection rate was 10 times normal. The exhaust gas temperature spread was used to indicate the performance of the nozzles in the fuel manifold.

A rejected fuel manifold with 98-hr operating time was disassembled for inspection and flow calibration. Deposits in the fuel manifold tubing ranged from a dark tan similar to a CRC color code No. 4 to a black coke. The nozzle spray patterns shown in Fig. 11 were extremely poor. Thirty-eight nozzles were plugged in the main fuel section, and the manifold had a 42% reduction in fuel flow. The main fuel screen and the swirl slots are illustrated schematically in Fig. 12. The heavy deposits observed on these screens and swirl slots are shown in Fig. 13. These plugged nozzles distorted the burner flame patterns, resulting in uneven turbine temperatures and warping of the turbine inlet guide vanes downstream, which adversely affects engine performance.

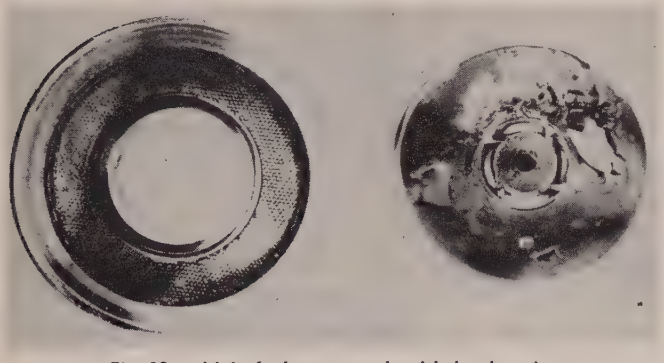


Fig. 13 — Main fuel screen and swirl slot deposits

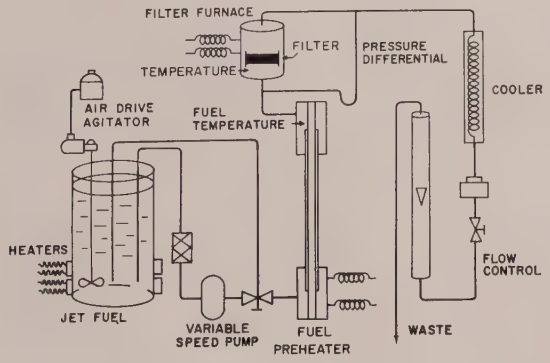


Fig. 14 — CFR Research Fuel Coker flow diagram

Navy Service Experience

The Navy Aeronautical Engine Laboratory conducted a survey of the thermal stability characteristics of a JP-5 fuel used at an East Coast Naval Air Station. The aircraft of principal interest flown by this squadron were the F8U and F4D equipped with J-57 powerplants.

Fuel samples taken from the truck servicing fuel directly to the aircraft were unstable in the CFR Fuel Coker at 400 F preheater and 500 F filter temperatures. The fuel was stable, however, at 300/500 F preheater and filter temperatures and a 6-lb/hr fuel rate. The preheater deposit rated code No. 1 and the filter ΔP was 5 in. of Hg after 300 min.

Six engines were inspected after operating times of 50–135 hr. The manifolds (dual-tube shielded) and fuel nozzles were free of deposits. There was some light discoloration in the 0–2 range of the CRC color code. The two nozzles nearest the fuel inlet in the nozzle cluster were the cleanest. The next two were slightly colored, and the last two were the darkest. The fuel nozzle screens and swirl-slot passages were clean, with a slight discoloration which followed the same pattern as the nozzles but was about one color code number darker. The overall performance of the engine fuel system was completely satisfactory.

High-Temperature Research Fuel Coker

The CFR Fuel Coker can be used effectively for the majority of today's jet fuels. Equipment similar to the CFR Fuel Coker is also needed in order to extend the temperature range above 500 F and evaluate hydrocarbon fuels for Mach 3 and higher speed aircraft.

A research fuel coker designed by the Equipment and Procedure Panel of the CRC Thermal Stability Group provides equipment capable of operating at temperatures considerably higher than those used in the standard CFR Fuel Coker. The equipment is still under test and may be subject to changes as the result of continued development with three prototype models.

A schematic flow diagram of the research coker is shown in Fig. 14. The preheater and filter sections are mounted vertically. Fuel is exposed to long heat transfer periods in a stainless-steel reservoir at temperature-time conditions similar to those in airframe wing tanks.

The fuel temperature in the 10-gal reservoir is

controlled by two 1500-w heaters mounted on the outside. The temperature is indicated by a thermocouple inside the reservoir just above the upper heater. The warmup time has been calibrated with a power input of 2500 w at the conditions shown in Table 7. A cam-operated controller programs the warmup rate and provides automatic control of reservoir temperature during operations.

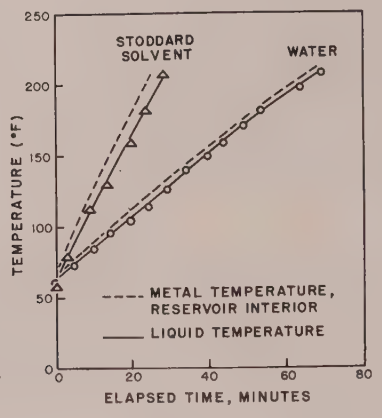


Fig. 15—CFR Research Fuel Coker fuel reservoir warmup

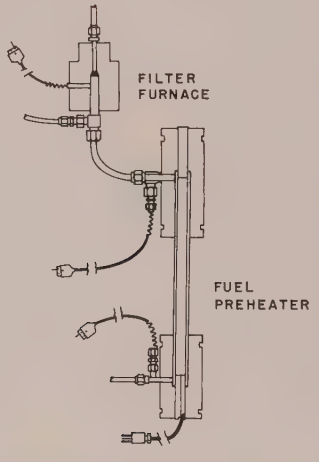


Fig. 16—CFR Research Fuel Coker

Table 7 — High-Temperature CFR Research Fuel Coker
Fuel Reservoir Warmup

	Calibrating Fluids	
	Water	Stoddard Solvent
Volume Tested, gal	6.1	6.1
Sample Weight, lb	50.9	39.6
Specific Heat, Btu/lb F	1.0	0.5
Power Input, w	2500	2500
Power Density, w/sq in.	6.8	6.8

Table 8 — High-Temperature CFR Research Coker
Operating Conditions

1. Preheater and Filter Temperatures	to 750 F
2. Fuel Flow Rate	2–8 lb/hr
3. Fuel Pressure	100–250 psig
4. Fuel Specific Gravity Range	0.7–1.2
5. Reservoir Temperature	Ambient to 250 F

Typical results, shown in Fig. 15, indicate that Stoddard solvent requires a shorter warmup than water because its specific heat is 50% lower.

A variable-speed-drive, positive-displacement pump is used with a regulator to control fuel pressure to 250 psig. Fuel flow is controlled manually by a needle valve near the downstream outlet of the system.

A bypass is provided for the test filter to permit any test to be continued in the event a limiting pressure differential is obtained prior to normal completion of the test. The bypass is a standard 3-valve unit developed for the CFR Fuel Coker to allow preheater deposits to be compared at the same test durations.

A comparison of preheater section dimensions for the CFR Fuel Coker and the Research Coker shows that the tube length and the annular fuel space between the inner and outer tubes are essentially the same, to provide equivalent Reynolds numbers. The Research Coker has a 2100-w electric heater and an enlarged inner tube with a thinner tube wall. The flow area and heat flux are higher than in the standard fuel coker. The preheater and filter furnace assemblies are shown in Fig. 16.

After passing through the preheater and filter assembly, the fuel is cooled by water heat exchangers. A filter located after the coolers prevents particles plugging the flow-adjusting valve. Indicator-controllers are used with voltage regulators and watt meters to govern the preheater and filter temperatures during warmup. The Research unit is capable of operating at the conditions summarized in Table 8.

The activities of the Equipment and Procedure Panel regarding the CRC Research Coker have been completed, and a separate High-Temperature Stability Group has been constituted to cooperate with the Air Force on future investigations.

Current programs related to super-stability evaluations of hydrocarbon fuels at temperatures above 500 F include those of the Navy Aeronautical Engine Laboratory, EPPI Precision Products, Inc. working cooperatively with CRC, and the Southwest Research Institute working under Air Force contract.

Summary

1. The laboratory techniques developed by the Coordinating Research Council for evaluating thermal stability at elevated temperatures have proved to be useful in studies of jet fuel applications.

2. It is possible to determine whether a jet fuel is thermally stable in the CFR Fuel Coker at a given temperature by means of a "pass-fail" criterion. The test clearly distinguishes stable fuels which are essential for safety and efficient engine utilization.

3. CFR Fuel Coker determinations are indicative of jet fuel performance in controlled flight tests and in military service.

4. The incidence of serious difficulties with fuel system deposits in Air Force service has been determined in a representative sampling of engines from heavy jet aircraft. The deposits caused by unstable JP-4 fuel were readily detected by the CFR Fuel Coker.

5. The performance of JP-5 fuel thermally stable at 300/500 F in the CFR Fuel Coker has been satisfactory in Navy service with no deposition difficulties in the engine fuel systems.

Soviet Agricultural Tractors and Equipment

Wayne H. Worthington
Heinrich Lanz, A. G.†

THIS PAPER presents an appraisal and evaluation of progress in the development of farm equipment in the USSR. Included in the discussion are present production wheel and diesel tractors, ranging from 18 to 54 hp; new models intended for 1959 production; and current wheel-type farm tractors and tracklaying farm tractors.

The Appendix lists specifications of tractor engines, starting engines, and the current tractors.

Material in this paper was gathered by the author during a 9000-mile trip through the Soviet Union. He was a member of the six-man Scientific and Engineering Team sponsored by the U. S. Department of Agriculture in 1958.*

LITTLE DIRECT comparison of Soviet agriculture may be made with that of the United States. Nor can one generalize with respect to the more specialized aspects of agricultural engineering and farm mechanization. Compared with similar developments in the United States, the various phases present unequal, and many times, contrary developments. More important than the actual stage of present development is the appraisal and evaluation of progress being made, and especially, the rate of progress. But in spite of great variations found among activities relating to research, education, and practice, the trends of good husbandry, efficient use of modern technology with respect to machinery and human resources, crop yields, and total production are definitely upward. The Government and the Communist Party place different priorities on facilities and set different goals than do agriculturalists in other countries.

Another significant factor relates to the cultural background of Soviet farm people, which differs significantly from that of the western world. Under strong Roman influence for generations, Western Europe fell heir to the Roman traditions of an independent farm management, and with it, the self-reliant farm family—an idea which in the course of history spread into what is now the United States and Canada. Until the liberation of some 15,000,000 serfs in 1861, Russian farm land was held largely by the Crown or members of a decaying feudal aristocracy. Following the liberation, agriculture remained primitive and advanced farming practices were infrequently undertaken, except for a few relatively large estates and more advanced communities.

Under the Czarist regime, the Russian Government suffered grave and continuing political difficulties and reverses, starting well before 1900 and continuing through World War I. These culminated in the Revolution which deposed the Czar, followed by the Communist Revolution and a bitter Civil War of extermination. Modern farming, as we understand the term, was not undertaken until the late 1920's. Despite the greatest of difficulties, and in a situation where everything needed doing at once, real progress was made. Then came World War II, the German invasion, and the "scorched earth" policy of the retreating Russian armies, bringing with them the destruction of the cities and devastation of farms throughout the greater part of the country's most productive sections. Over great areas, famine stalked. Livestock was slaughtered and the land laid waste. The flower of Soviet young manhood was wiped out, leaving behind only women, children, and old men. Following hostilities, industry and the transportation system both had to be rebuilt before farms could be supplied with essential equipment, fertilizers and other materials necessary to rebuilding the soil and restoring the land to production. Concurrently, the livestock population had to be re-established and habitations for millions provided in both cities and farms.

Soviet statistics are consistently inconsistent and in some cases, intentionally misleading. For example, signs displayed at the 1959 Soviet Cultural Exhibition in New York gave the following data re-

† Formerly with John Deere Tractor and Engineering Center.
* Paper presented at SAE National Farm, Construction, and Industrial Machinery Meeting, Milwaukee, Sept. 16, 1959.

garding farm acreage:

Collective Farms		State Farms	
Number	69,320	Number	6000
Average Acreage	10,193	Average Acreage	60,039
Total Acreage	706,578,760	Total Acreage	306,234,000

As of July 27, 1959, the Soviet Embassy in Washington gave the number of collective and state farms as 78,200 and 5900, respectively. Current acreage in grain production is given as:

All Crops	483,400,000 acres
All Grain	309,300,000 acres
Wheat	164,700,000 acres
Corn	48,700,000 acres

It is believed that the above does not include the millions of small land plots, ranging from ⅓ to 2 acres, rented to each farm family on all of the state and colective farms for its own use. It is now less than two years since the Machine Tractor Stations throughout the country disposed of 68% of all tractors (406,200 units) and 43% of all combines (136,900 units) to the collective farms. Yet according to official statistics applicable to the same period, 6% of all grain (18,560,000 acres) is harvested manually by sickle and scythes and 7,186,000 acres by horsedrawn reapers.

This great forced consolidation of farm land into collective and state farms has resulted in a corresponding concentration of management and technical skills and responsibilities. In other ways, it facilitates government planning and the enforcement of controls, and makes possible prompt and widespread establishment of new practices. Yet many such practices have been ill-suited to particular areas and great confusion has resulted.

Much of the mechanization effected on farms appears to have been more with the idea of increasing food production rather than effecting a reduction in human toil and the freeing of manpower for industry. Since the death of Joseph Stalin in 1953, a severe labor shortage has developed and there has been a greater emphasis placed upon the labor-saving aspects of farm mechanization.

One cannot disregard the heterogenous nature of the Soviet Union, with its many lands and ethnic groups, arranged into 15 so-called "Autonomous Republics." Of these, the Russian Republic is by far the largest and probably plays a dominant role in the political life of the Union. As a result of the 82 different languages spoken throughout the country, great language barriers exist, which no effort is being spared to break down. But no man-made effort can reduce the great differences existing in the soils, climates, vegetation, people, ideas, customs, cultures, and politics throughout the vast reaches of the entire land.

Based on statements made by the Deputy Minister of Agriculture responsible for Farm Mechanization, there were a total of 597,350 individual tractors in the Soviet Union "at the end of 1957." Taking the total power of all Soviet tractors at 22,500,000 hp (1,700,000 tractor units of 15 hp, as advertised at the New York Exhibition), the average Soviet farm tractor as of that time had a engine developing 37.7 crankshaft hp. From the most reliable information available, some of which was verified by factory production boards, 1958 tractor production (including both tracklaying and wheel types) was at the daily rate of 608 units, for a total annual production of approximately 180,000 units. Some of these are known to have been exported to other Communist countries and to India. Not counting the specially built logging tractors and those of over 80 hp, only a few of which were seen on farms, the power of all Soviet farm tractors built in 1958 averaged out to 39.1 hp, in terms of metric horsepower at the engine flywheel. All Soviet farm tractors are currently powered with 4-stroke, solid injection, water-coooled, diesel engines of the vertical in-line type. In contrast to American-built tractors, in which their power at the belt is reported as that observed at the dynamometer with no allowance for transmission or belt losses, Soviet tractors are rated in terms of net engine metric flywheel power. Power losses between engine and dynamometer can be considerable. For instance, the 1956 Model Byelaruss MTZ-2 tractor had a published rating of 37 hp. Test Report No. 253 of the Finnish Machinery Testing Station (Valt Moatalousk Tutkimusl Helsinki) officially reports the performance of this tractor as follows:

Maximum belt hp (corrected	33.9
Specific fuel consumption (hp hr)	0.5148 lb
Drawbar hp — observed	
Asphalt track	30.5
Cultivated soil	20.9

In considering the performance of Soviet tractors, one must not lose sight of the fact that their power rating is in terms of metric horsepower. One metric horsepower is 75 kg-m (542.48 ft-lb) per sec, or 0.9863 hp.

In the face of a severe domestic shortage of power farming equipment, the Soviet Union exports tractors and implements as a matter of political expediency. Export prices in no way reflect manufacturing costs as we of the Western World know them. For instance, in September, 1958, Soviet trade officials were offering their best wheel tractors, with rubber tires, delivered at Antwerp, Belgium, at prices of $0.188 U. S.)–$0.21 per lb. As a further inducement, all tractors and stocks of repair parts were offered on consignment with no down payment.

Farm tractors in the Soviet Union fall into three general classses:

1. Older tractors, built before and in the years immediately following World War II, the design of which closely follows that of their U. S. prototypes. In the case of wheel tractors, their manufacture was abandoned as soon as new designs could be placed in production following World War II. With the exception of the large crawler tractors built in western Siberia, which follow the design of an early American tracklayer, these earlier tractors were powered with carburetor-type engines and burned either gasoline or distillate.

2. Present production wheel and diesel tractors, of 18–54 hp, which are "genuine Soviet originals." These constitute the greater number of tractors now in use. Their power-to-weight ratio is low, specific fuel consumption high, and general design heavy handed.

3. New models intended for 1959 production or later. These incorporate many distinctive and modern improvements, including increased power and power-to-weight ratio, lower specific fuel consumption, travel speeds better suited to the operation of higher speed implements, and greater operator's convenience and comfort.

L. A. Korbutt, Deputy Minister of Agriculture, stated that as of 1953, 52–53% of all Soviet farm

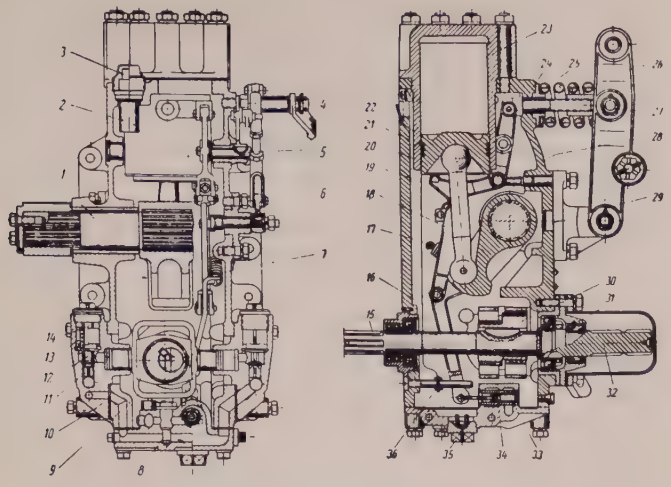


Fig. 1 — Hydraulic implement-lifting and draft-regulating cylinders with servo controls

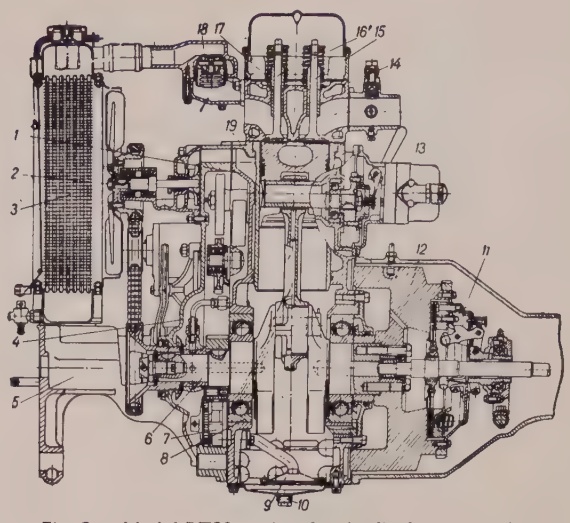


Fig. 2 — Model DT20 engine, longitudinal cross-section

tractors were tracklayers. Due to the expansion of row crop acreage, increasing mechanization of transportation, and lower first cost, it is planned to increase the number of wheel tractors to 70% of the total within 5 years. We also learned that:

1. Tractors are sold to the Tractor Repair Stations for delivery to collective farms at factory production cost plus freight.

2. The average "mark up" or delivery charge, made by the Machine Repair Stations is 10–15%. The operators of these stations in the remote reaches of Siberia, where all-out mechanization is imperative, state that they make no delivery charge and that any deficit in operating expenses is chargeable to the State budget.

3. The value of all repair parts manufactured is 60% of the value of the tractors produced.

Specifications of Soviet tractors, gathered from many sources, are presented in the Appendix.

For several years following World War II, the "Universal" wheel-type tractor, and more recently, "Byelaruss" wheel type and the KDP-35 tracklaying cultivating tractors were equipped with a 3-point implement hitch and a Model HC-37 hydraulic implement lifting and draft control system. A cross-section view of the latter is shown in Fig 1. Hydraulic power was generated by a 4-plunger pump, with bore and stroke of 1.18 in. and 0.787 in., respectively, operating at speeds of 520–660 rpm. Control of pump delivery is by throttling of the pump inlet. The single-acting lift cylinder, with bore and stroke of 3.94 in. and 5.51 in., respectively, is positioned vertically at the rear and the lift time ranges from 3 to 3.75 sec. This compares with the ASAE-SAE standard lifting time of 1.5–2.0 sec.

In 1957, the use of the automatic draft control feature was discontinued, because Soviet agronomists, supported by the Scientific Council of the Ministry of Agriculture, felt that the variations in implement working depth, which characterize this type of control, could not be tolerated. At present, integral tillage implements attaching to all 3-point tractor hitches are either allowed to "float," or their working depth is gaged from the tractor by means of a hydraulic cylinder with the working depth controlled by the driver. In 1958, in an effort to effect

Table 1 — Specifications of Complete Hydraulic Control Systems

Tractor Hp	14–16	28–32	40	55–70	80–100
Pump Output, gpm	4.33	10.6	12.9	15.9–19.8	40
Cylinder Diameter, in.	2.95	3.54	3.94	4.35	4.73
Stroke in Force,					
Maximum, lb	8800	13,200	16,500	19,800	24,200
Pressure, psi	12,880	13,400	13,500	13,750	13,800

standardization of components, the manufacture of complete hydraulic control systems, for use as original equipment by a number of tractor plants, was undertaken by one specialized factory in Moscow. Features characterizing these new systems are:

1. All pumps are of the gear type and are designed to be driven from the engine timing gear train.

2. Cylinders are both single and double acting, with chrome-plated piston rods.

3. The system is of the open-center type. Valves may be "stacked" in series for the operation of two or more cylinders.

4. Operation of the control valve used with integral implements is such that when the manual operating level is placed in its raising position, it is retained by means of a detent. When the cylinder reaches the end of its lift stroke, hydraulic force is applied to overcome the detent and force the valve into the "hold" position.

5. A "float" position is provided.

Specifications of these systems are shown in Table 1. The values of maximum cylinder force, applying to the four largest cylinders, represent relief valve opening pressures averaging 1360 psi (960 kg/sq cm) and are comparable with the best current United States practices applying to the use of gear pumps. Study of this standardizing effort indicates no consideration has been given to the interchangeable use of cylinders of various capacities on any and all tractor-drawn implements, as is the case with our comparable ASAE-SAE standards.

Current Soviet Wheel-Type Farm Tractors

Model DT–20, General Purpose Vegetable and Garden Tractor — Manufactured at rate of 80 per day

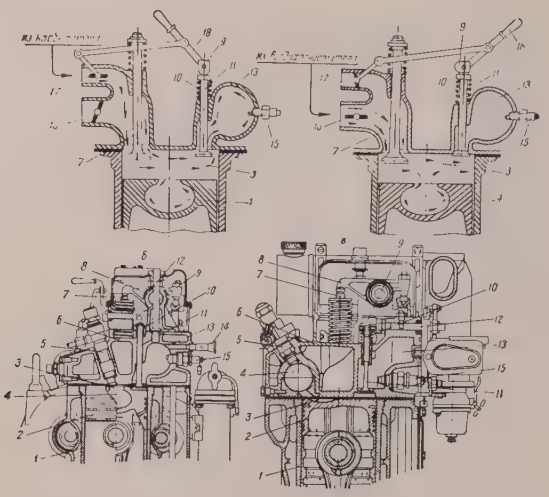


Fig. 3 — Auxiliary combustion chambers for gasoline starting (upper left, upper right, and lower left — Model DT20; lower right — Models DT24 and DT28)

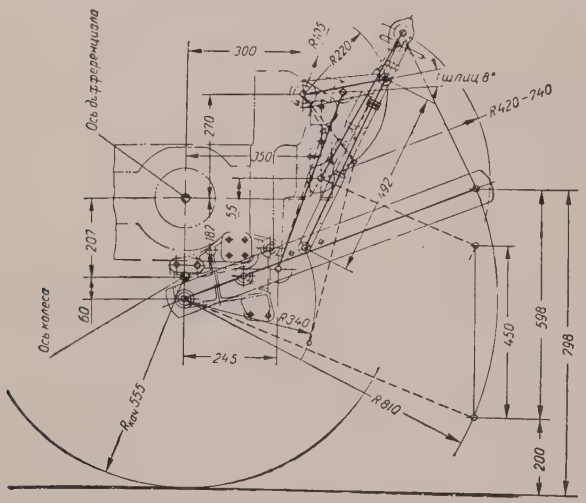


Fig. 4 — Model DT20 tractor, geometry of 3-point implement linkage

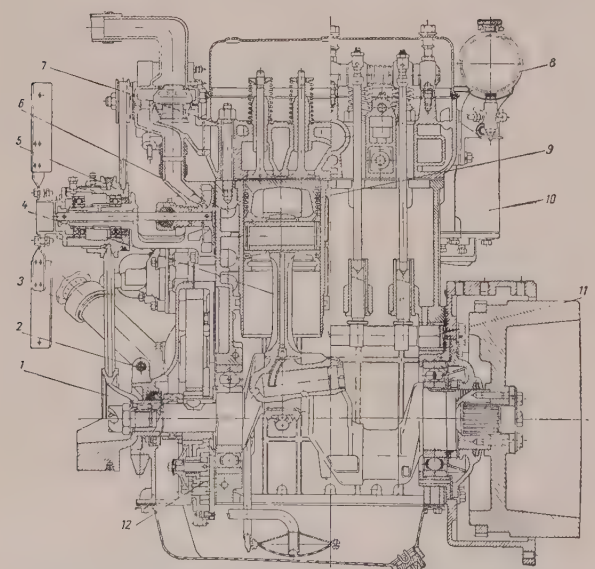


Fig. 5 — Model DT28 tractor, longitudinal section of engine

by Kharkov Tractor Works.

This small single-cylinder tractor is intended for vegetable and garden operations, where only limited power is required. Manufacture started in 1955 as the Model DT-14. A longitudinal section of the engine is shown in Fig. 2. Starting is effected by opening an auxiliary combustion chamber, which reduces the compression ratio from 15.0/1 to 5.6/1 to permit gasoline operation and cranking with 12-volt starting motor. The diesel combustion chamber with the auxiliary combustion chamber is shown in Fig. 3. The engine is provided with a Lancaster-type balancer, which employs two counter rotating shafts with balance weights, operating at crankshaft speed. This balances out vertical primary forces except those due to the finite length of the connecting rod. The geometry of the 3-point implement linkage is shown in Fig. 4.

This versatile little tractor is provided in four modifications, obtained by rotating the final drive housings with respect to the main axle housings, as shown in Table 2.

Model DSSH–14 Self-Propelled Chassis — Manufactured at the rate of 20 units daily by the Kharkov Tractor Assembly Works.

This small implement carrier has recently been put into production at the Kharkov Tractor Assembly Works. It is powered with the same single-cylinder engine as used in the DT–20. Like its German-built prototype, the Lanz "Alldog," it con-

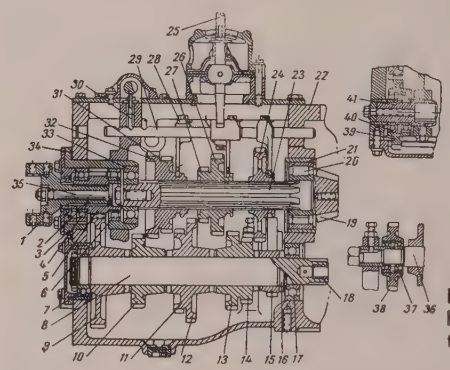


Fig. 6 — Model DT24, longitudinal section through 5-speed transmission

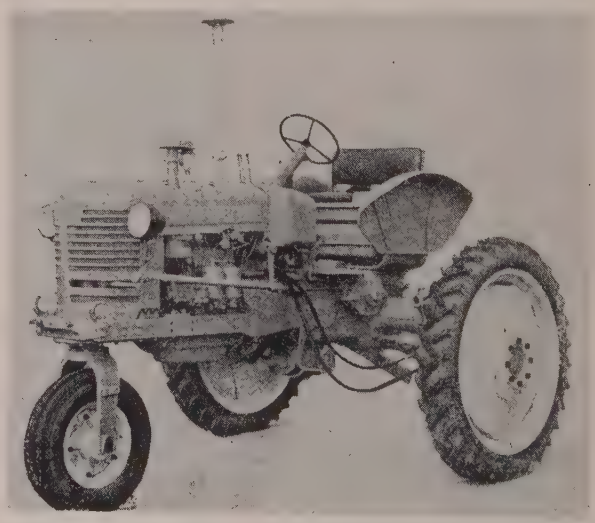


Fig. 7 — Model DT24-3, high-clearance tricycle cultivation tractor

sists of an engine and transmission unit located at the rear of a rectangular tubular frame. A wide variety of integral implements may be attached between the front and rear axles. The principal specifications of this tractor are listed in Table 3.

Three power outlets are provided:

1. Front-mounted belt pulley.
2. Independently driven power take-off with a speed of 533 rpm.
3. Powershaft with speed proportional to ground travel, operating over a speed range of 325–1695 rpm in first to fifth gear.

The tractor is regularly equipped with a hydraulic lift, powered with a 2-cyl plunger-type pump, 1.18-in. bore × 0.787-in. stroke, which delivers 4.04 gpm at a working pressure of 1138 psi. Two small single-acting cylinders are supplied, each with a bore of 1.42 in. and a working stroke of 9.84 in.

Implements available at present include:

2-bottom, one-way plow (10-in. maximum working depth)
6- and 8-row vegetable planter
4-row vegetable cultivator with fertilizer attachment
6-row cotton cultivator, arranged for rows spread at 24 in.
6-row cotton cultivator, arranged for rows spread at 16 in.
10-ft buck rack (front mounted)
1-row potato digger
Spray tank with pump and booms
7-ft mower cutter bar
Dump box with struck capacity of 0.9 cu yd
3-row sugar beet lifter
Crop duster with fan, 32.8 ft
10-ft fertilizer distributor
General-purpose wagon box, arranged for dumping to either side or to front

It was stated that it is expected to eventually provide a total of 22 different integral implements.

Table 2 — Modifications of Model DT-20 Garden Tractor

Modification	Position of Wheel Center with Respect to Axle Housing	Wheelbase, in.	Hood Height, in.	Ground Clearance, in.
A	Directly below	65.0	56.5	20.25[a]
B	90 deg forward	56.2	48.5	12.12[b]
C	90 deg rear	72.5	48.5	12.12[b]
D	45 deg rear	70.1	56.5	16.1[a]

[a] Clearance under transmission case.
[b] Clearance under front axle.

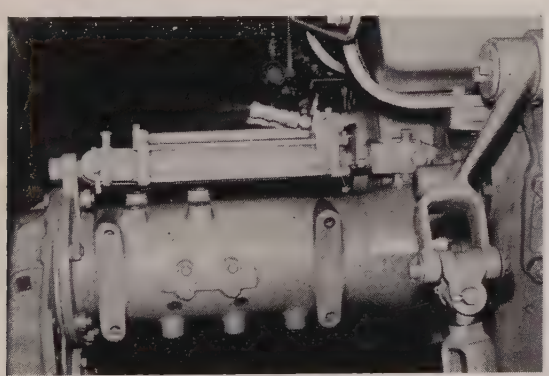


Fig. 8 — Model DT24-3 high-clearance tractor with hydraulic rear wheel tread adjuster

This appears to be a very ambitious program, considering the limited power of the tractor.

Model DT–24, General Purpose Tractor — Manufactured at the rate of 60 per day by the Vladimir Tractor Works.

This small general-purpose tractor is powered with a 2-cylinder, 4-stroke diesel engine, a longitudinal section of which is shown in Fig. 5. Like the engine used in the Model DT–20 tractor, this engine is provided with an auxiliary combustion chamber which can be made operative to reduce to the compression ratio from 14.5/1 to 5.5–6.0/1 for starting as a spark ignition engine with gasoline. This auxiliary combustion chamber is shown in the lower right of Fig. 2. A longitudinal section through the 5-speed sliding gear transmission is shown in Fig. 6.

This basic tractor is used to provide three modifications:

DT-24-1	Tricycle Model, with double front wheels
DT-24-2	Standard tread 4-wheel model
DT-24-3	High-clearance tricycle model, with rear drop axles and single front wheel

With the first two modifications, the rear wheels are adjusted laterally by shifting rims with respect to wheel disks and reversing the disks on the axles. The high-clearance Model DT–24–3 is shown in Fig. 7. Adjustment of rear wheel tread is made by means of telescoping rear axle quills, which are forced to slide axially by means of a hydraulic cylinder. This construction is shown in Fig. 8.

It will be seen that the larger external quills are flanged directly to the sides of the transmission case. Internal quills, having the rear axle drop housings attached to their respective outer ends, slide laterally within the outer quills to effect the rear wheel tread adjustment.

The relatively large 2-cyl engine (4.91 × 4.91 in.) is provided with a decompression gear, but is hand cranked. The 3-point hitches provided for all modifications of these tractors mount a wide variety of integral and semi-integral implements and a drawbar attaching to the lower draft links may be used to pull available drawn implements.

Model MTZ–5 Byelaruss Wheel-Type Tractor — Basic model manufactured at a rate of 100 units daily by the Minsk Tractor Works and at a rate of 60 units daily at the Dniepropetrovsk Machine

Table 3 — Specifications of Model DSSH-14 Tractor

Wheelbase, in.	98.5
Tread	
Drive wheels, in.	47.24–70.8
Front wheels, in.	47.24–59.06
(Front & rear wheels adjustable in increments of 3.94 in.)	
Ground Clearance, in.	22.0
Operating Weight, lb	2997
Tire Sizes	
Front	5.50 × 16
Rear	8 × 32
Travel Speeds, mph	
Creeper	0.802
1st	2.07
2nd	2.67
3rd	3.38
4th	4.21
5th	10.7
Reverse	1.09
Minimum Rear Wheel Turning Radius, ft	10.08

Works.

This is the largest wheel-type general-purpose tractor now built in the Soviet Union. It is powered with a 4-cyl, 4-stroke in-line diesel engine, having a Ricardo-type spherical combustion chamber. It is regularly equipped with a 3-point hitch and hydraulic lift made to the latest Soviet standards. Electric starting motor or single-cylinder gasoline starting engine are optionally supplied.

The most popular Byelaruss modification is the 4-wheel general-purpose high-clearance type, with adjustable front and rear treads and 3-point implement hitch. Soviet farmers have never accepted the tricycle-type row crop tractor, with either one or two front wheels, which is so widely used in this country. The tricycle tractor, with the front-mounted integral tool frames carrying tillage and cultivating tools visible from the operator's seat, is strictly a one-man machine. All of the integral 3-point mounted implements displayed at the Minsk Tractor Works were provided with manually adjusted gage wheels to maintain implement working depth. Many of the Soviet rear-mounted 3-point cultivators and implements require a second operator on the implement to "steer" them. This wasteful use of farm labor is the accepted Soviet way of doing things and is a major factor contributing to the low productivity per worker.

In addition to the Byelaruss wide-tread and triangle modifications, a 4-wheel drive modification is provided which greatly improves performance, especially locomotion, in swampy and extreme muddy conditions. On the 4-wheel drive modification, the weight of the front end of the chassis is supported by a lateral leaf spring, with its ends shackled to the front axle. A front view of this tractor is shown in Fig. 9. As compared with the wide-tread 4-wheel general-purpose model, the weight on the front wheels of this 4-wheel drive modification reflects only the difference in weight between the conventional axle and the 4-four wheel drive axle with a part of the weight of the propeller shaft. As with all Byelaruss tractors, a differential lock is available for the rear axle, but there is no differential action between the drives for the front and rear axles. No performance data is available for this tractor.

The only official information regarding the performance of comparable 4-wheel drive tractors (built by modifying conventional rear-wheel drive tractors by adding a front-wheel drive "attachment") is that provided by the official Swedish tests of the 2-wheel and 4-wheel drive MAN (German built) tractors. Table 4 shows the performance data of a single tractor, operated with rear-wheel and 4-wheel drive. As a matter of information, and for the evaluation of the performance of such tractors, drawbar performance and wheel slippage data are presented.

In summary, under test track conditions:

1. The maximum drawbar pull of the 4-wheel modification is greater than that of the comparable 2-wheel drive.

2. In second and third gear, the drawbar pull and drive wheel slippage are less with 4-wheel drive.

3. Maximum drawbar power is greater with the 4-wheel drive tractor.

Tests of this kind, which are comparable with our Official Nebraska Tests, do not truly show the benefits in traction, locomotion, and mobility which characterize the superior performance of 4-wheel drive tractors under swampy and other poor traction conditions.

Table 4 — Performance Data, MAN 25-hp Plowing Tractor[a]
Comparative Drawbar Tests[b]

	2-Wheel Drive			4-Wheel Drive		
Gear	1st	2nd	3rd	1st	2nd	3rd
Maximum Drawbar Power, hp[a]	14.0	18.5	18.8	17.5	19.6	18.8
Drawbar Pull, lb	2970	2175	1410	3655	2085(2)	1362(2)
Travel Speed, mph	1.74	3.12	4.92	1.74	3.49	5.11
Wheel Slip, %	24	14	6	21	5	3
Ratio $\frac{\text{Drawbar, hp}}{\text{Belt hp}}$	—	0.733	0.744	0.696	0.778	0.744

[a] Metric hp. One metric horsepower = 0.9862 hp in English units.
[b] Data from Official Swedish Tractor Test Nos. 942 and 943. Issued by the State Agricultural Machinery Testing Institute. Same tractor is used for both tests.

Weights, lb (test weight with driver)	
Front	1980
Rear	3365
Total	5345

Note: In second and third gear, the drawbar pull in 4-wheel drive is slightly lower than in rear-wheel drive. This is because of less wheel slip and a consequently higher travel speed in 4-wheel drive.

Current Soviet Tracklaying Farm Tractors

Aside from the "Stalinetz" Model S-100 crawler, which is essentially an "industrial" tractor rather than "farm" tractor, there are two basic Soviet tracklaying farm tractors in current production — the KDP-38 and the DT-54, each with several modifications.

All Soviet tracklaying farm tractors (other than the S-100) are equipped with tracks having track links and shoe cast in one piece of austenitic steel (12%–12% manganese). As this material is too hard to machine, all holes for the track link pins are cored, and the tracks are assembled in their "as-cast" condition. The track pins are of carburized cold-drawn steel, with cross-hole drilled pin holes at the ends to accommodate wire retaining rings. The construction of the track link is shown in Fig. 10. Each track of the DT-54 tracklayer comprises 42 individual cast links, weighing a total of 704 lb or approximately 17 lb per link. Track life varies greatly from 600–800 hr under severe sandy conditions upward to 3000 hr working in black loam. The Managing Director of the Kharkov Tractor Works stated that the "manufacturing cost" of track castings, made of 13% manganese steel was $6.82 per cwt. This totals approximately $48.00 per track. Track cost to the users is further reduced by the re-purchase of scrap tracks by the Tractor Works at $2.28 per cwt, FOB factory. It will be seen that the hourly track cost is very low, even in terms of wheat, which in most localities delivers to the elevators of the State Grain Trust at prices averaging $2.13 per bushel.

The various types of tracks suspension used on Russian tracklayers are shown on Fig. 11.

Rigid tracks are currently used only in an experimental way on a small orchard crawler tractor developed (but not yet placed in production) by the Kharkov Tractor Works. Semi-rigid tracks, shown

Fig. 9 — Byelaruss 4-wheel drive tractor

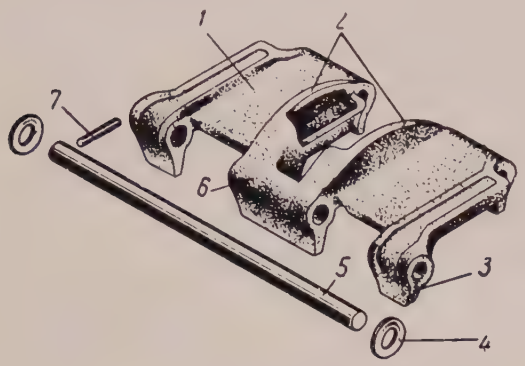


Fig. 10 — Cast manganese steel track shoe with pin

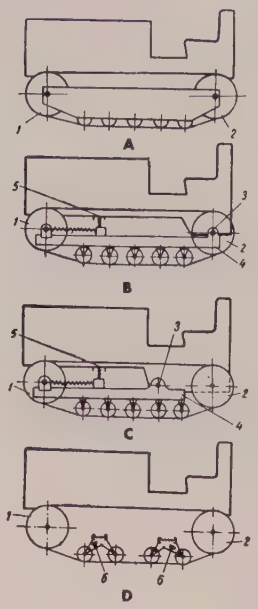


Fig. 11 — Types of track suspensions (A — rigid, B and C — semi-rigid, D — articulated bogies)

schematically in Fig. 11B, are used on large Model S100 industrial tractor, while Fig. 11C typifies the mounting on the general-purpose cultivating tractors built at Lipetz. Articulated bogies, pivotally mounted on a rigid tractor frame, as shown in Fig. 11D are used on the widely produced TD-54 and some of the newer prototypes described later.

"Kirovetz" Models KDP-35 and KDP-35-2 — Built by the Lipetz Tractor Works at the rate of 55 units daily (to be replaced by KDP-38, now in pilot production).

This basic utility tractor is built in two modifications, intended for distinctly different types of service. Model KDP-35 is intended for general-purpose work (although the term does not have the same broad significance as employed in the United States). In many areas, only drawn implements are used. Ground clearance under the transmission case and rear axle housings is 25.2 in. Four-roller tracks with a wheelbase of 67.2 in. are used, with cast track shoes 7.87 in. and 11.4 in. wide optionally available. The basic 5-speed clash gear transmission provides five forward travel speeds from 2.37 to 5.65 mph and two creeper speeds of 0.51 and 1.58 mph. A longitudinal section of this tractor is shown in Fig. 12.

The KDP-35-2 modification is a narrow high-clearance tractor intended especially for use in vineyards and for cultivating narrowly spaced row crops. Ground clearance under the transmission and rear axle is 31.5 in. This tractor is provided with the same basic 5-speed clash gear transmission used on the KDP-35 and, in addition, has optionally available a total of eight creeper speeds ranging from

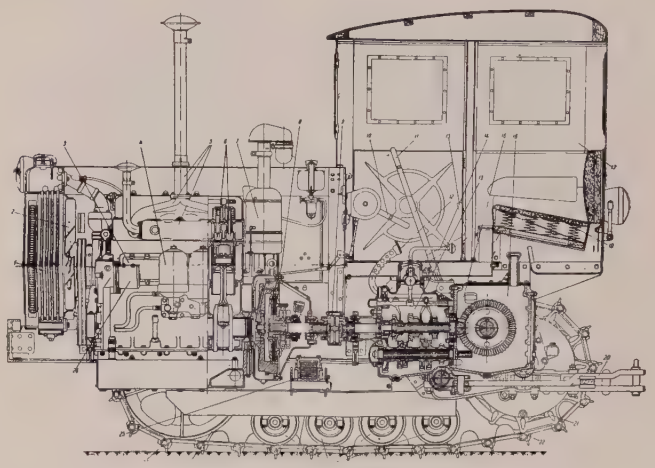


Fig. 12 — Model KDP-35, general-purpose track layer

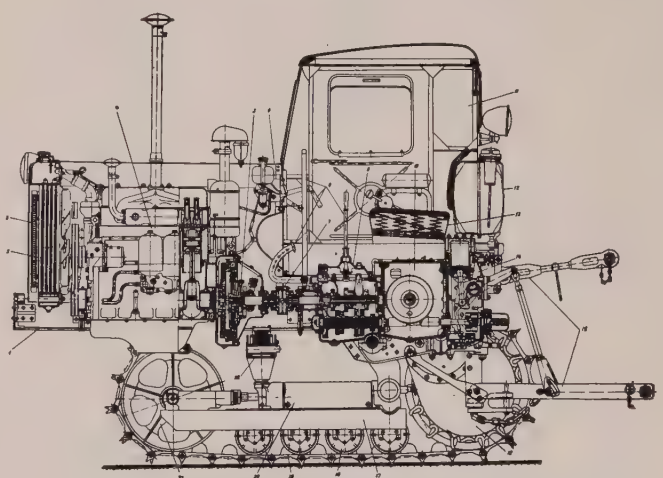


Fig. 13 — Model KDP-35-2, high-clearance cultivating tractor, longitudinal section

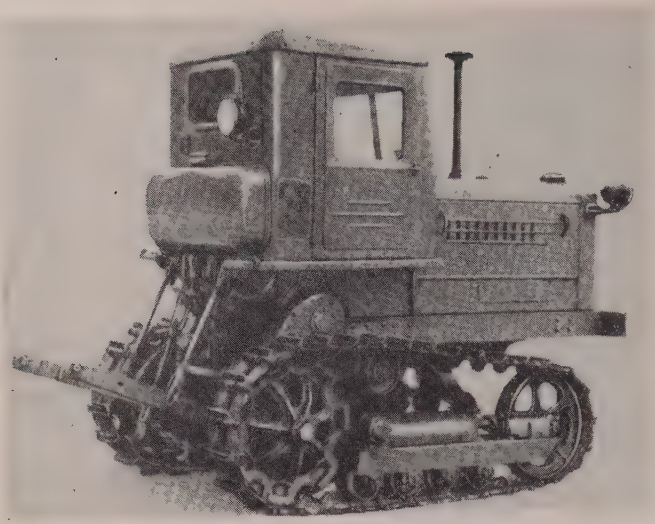


Fig. 14 — Model KDP-38-2, high-clearance cultivating tractor

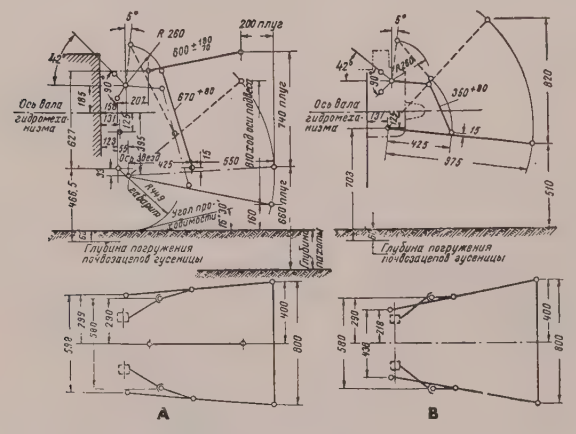


Fig. 15 — Geometry of 3-point implement hitches (A — Model KDP-35, B — Model KDP-35-2)

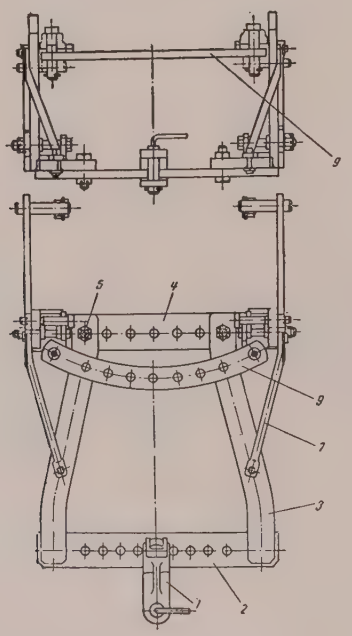


Fig. 16 — Model KDP-35 rigid drawbar

0.52 to 1.74 mph. Formerly, this tractor was provided with 5-roller tracks with a wheelbase of 82 in. To effect shorter turning on narrow head lands, the 4-roller tracks used on the KDP-35 modification with a wheelbase of 67.2 in. have replaced the 5-roller tracks. Track shoes only 6.29 in. wide hold down overall width and permit of operating in very close row spacing. A longitudinal section of this tractor is shown in Fig. 3. It will be noted that this shows the automatic hydraulic draft control system, which is no longer in production.

At one of the Research Centers visited, there was shown a KDP-38 tractor with 4-roller tracks and a wheelbase of 67.2. A view of this tractor with 3-point implement hitch is shown in Fig. 14.

Both modifications are powered with the same

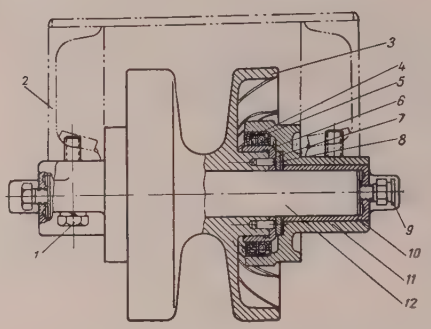


Fig. 17 — Track roller with plain bushings used on KDP-35 and KDP-35-2

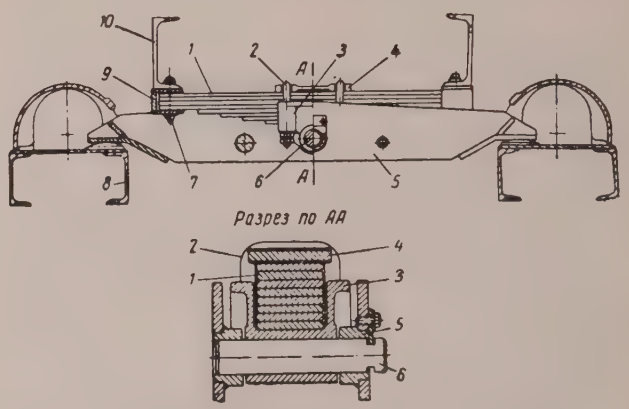


Fig. 18 — Models KDP-35 and KDP-35-2 lateral track equalizing spring

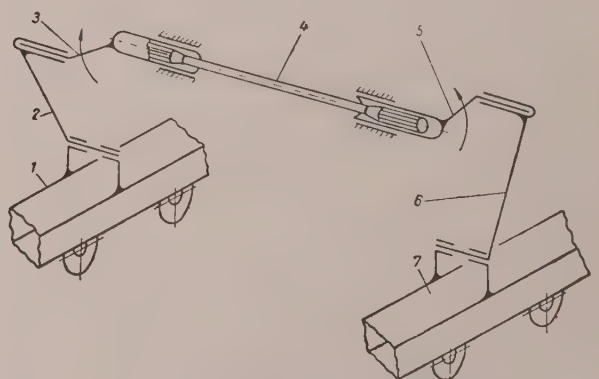


Fig. 19 — Preproduction torsion bar suspension for semi-rigid tracks

4-cyl engine as used on the Byelaruss MTZ-5 wheel tractor, the power of which is currently being increased to a maximum of 45 bhp. Both are optionally provided with 3-point implement hitches which accommodate a large variety of rear-mounted tillage and cultivating tools. Fig. 15 shows the geometry of the hitches — A shows the KDP-35 and B the KDP-35-2. The rigid drawbar regularly supplied for the Model KDP-35 is shown in Fig. 16. Both are provided with semi-rigid tracks, with the track rollers carried on plain bushings in housings rigidly attached to the track frame. This track roller construction is shown in Fig. 17. The front end of the frame of Models KDP-35 and KDP-35-2 is supported by a lateral equalizing spring, as shown in Fig. 18. Lip-type oil and dirt seals, such as those shown, are not considered adequate for use in severe dusty conditions in this country. The front end of the frame of the experimental KDP-38 (Fig. 14) is flexibly supported by a torsion bar. This construction is shown diagrammatically in Fig. 19.

The rapidly increasing power requirements necessary for Soviet farm operations, resulting from the use of wider and heavier integral tillage tools and cultivators and higher travel speeds, can be expected to be reflected by extensive changes in the design of these tractors.

DT-54 Tracklayer Tractor — Manufactured at the rate of 80 per day at the Kharkov Tractor Works, 110 per day at the Stalingrad Tractor Works, and 40 per day at the Altai (Rubtsov) Tractor Works.

The DT-54 is the heaviest and most powerful of the strictly agricultural crawler tractors currently in production. This tractor has reached a high degree of development, as it has been in production since 1949, and is acclaimed the most reliable and trouble-free of all Soviet tractors. The principal external dimensions are shown in Fig. 20. It has an independent frame as shown in Fig. 21. The tracks employ cast manganese steel shoes, with all weight supported on four sets of flexibly sprung bogies, the construction of which is shown in Figs. 22 and 23. Fig. 24 schematically shows the geometry of the bogie linkage. The track idlers are shown in Fig. 25. It will be noted that these idlers are mounted on tapered roller bearings and provided with spring-loaded face seals of excellent design.

Power is provided by a 4-cyl, 4-stroke diesel engine

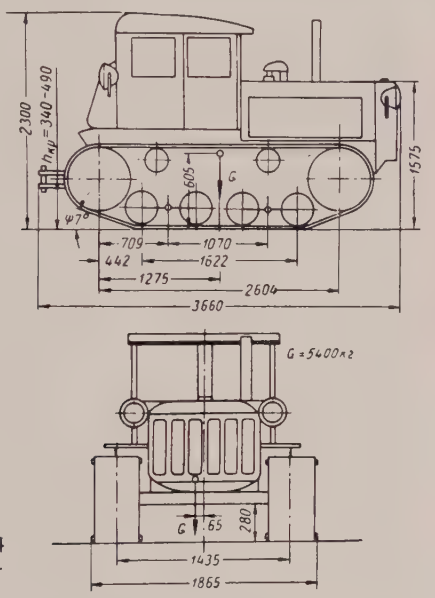


Fig. 20 — Model DT54 principal external dimensions (mm)

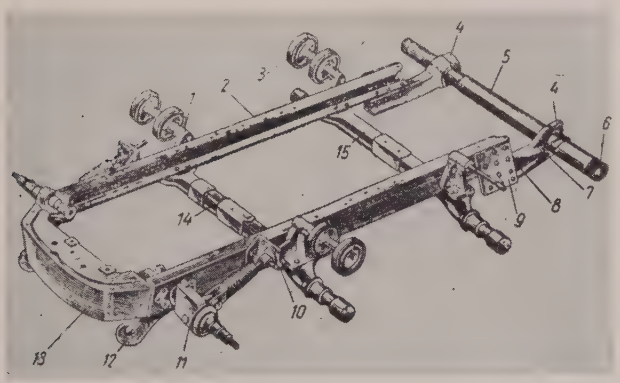


Fig. 21 — Model DT54, built-up frame

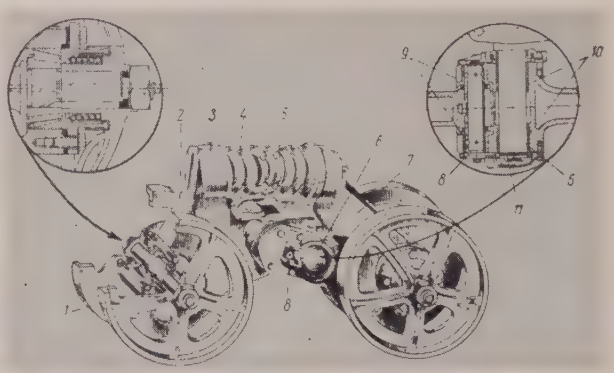


Fig. 22 — Model DT54 articulating track bogie

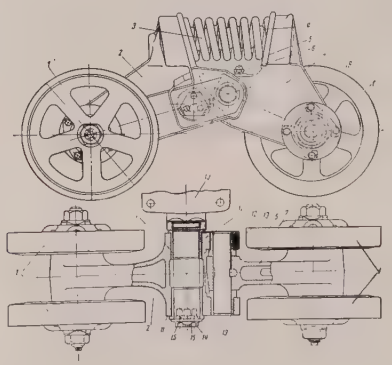


Fig. 23 — Model DT54, details of articulating track bogie

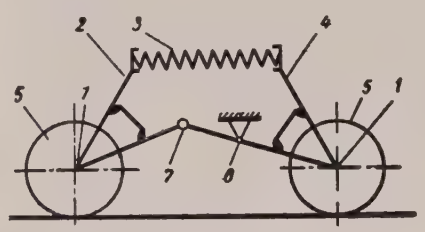


Fig. 24 — Model DT54 tractor, geometry of bogie linkage

(4.92 in. × 5.98 in.) operating at 1300 rpm. This engine is of heavy, slow-speed design, with low specific output and relatively high specific fuel consumption. On the other hand, the design is most conservative, due in part to the fact that the quality of the materials available during the period of its design and development were decidedly inferior to those available today. All crankshaft journals are induction hardened, and all main and conecting rod bearings are of the thin-shell steel-backed type, lined with leaded bronze.

This engine has a number of novel and interesting features:

1. The combination chamber is cylindrical, as shown in Fig. 26B. The entire combustion system has been the object of intense study by the State Scientific Research Institute for Farm Mechanization and Electrification, with the result that there has been developed a more conventional spherical swirl chamber with related depressions in the piston head (Fig. 26A). At the same time, the included cone angle of the fuel spray has been increased from 15 to 40 deg. It is claimed that the specific fuel consumption for the bare engine, without accessories, has been reduced from 0.493 to 0.402 lb/hp-hr, which is comparable with performance generally obtained with the better swirl-chamber-type engines. It is planned to place the new combustion system in production at an early date.

2. Great care is given to cleaning the crankcase oil, with the result that the factory recommendation for the service period between oil changes has been increased from 100 hr (when treated paper oil filter elements were used) to the present 250 hr. Three-step oil filtration is used, namely:

a. Edge wound wire screen, spaced at 0.0012–0.0016 in.
b. Reaction-driven centrifuge. A dirt separation bowl turns at 5000–6000 rpm, separating dirt from oil, collecting it in the bowl. This device is shown in cross-section in Fig. 27.
c. Drilled crankpins with submerged oil leads, as shown in Fig. 28. Due to centrifugal action, dirt separates from oil and accumulates in drilled pin. Dirt removed on occasion of general repairs at intervals of 1500–2000 hr.

3. In addition to a 10-hp single-cylinder, 2-stroke starting engine which heats the main engine lubricating oil during the cranking period, the intake air is heated and an electric glow plug is used to heat the combustion chamber. This engine is considered an "extra good starter," at temperatures as low as –40 F.

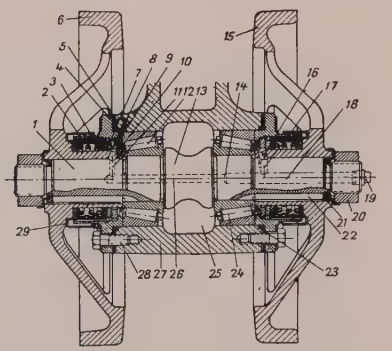


Fig. 25 — Model DT54 tractor, details of track idlers

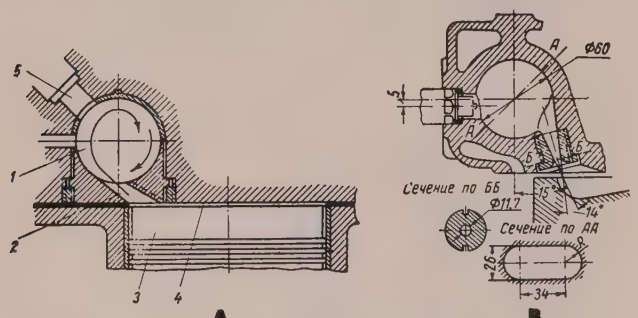


Fig. 26 — Model D54 engine, old and new combustion chambers

Two major modifications of the basic DT-54 tractor have been made by the Stalingrad Tractor Works to adapt them for special requirements. The first of these, introduced in 1955, is the Model DT-55, which is built for swamp and marshland operation. The modifications consist of increasing track width from 15.35 to 21 in., and pitch from 6.85 to 7.16 in. The diameter of the track idler wheel is increased to approximately 28 in., which extends ground contact from 63 to 88.5 in. The total weight increases from 10710 to 12750 lb, but the unit soil pressure is reduced from 7 to 3.58 psi. Creeper speeds of 0.30 and 0.39 mph are regularly provided, and optionally available are the six additionally creeper speeds (0.45 through 1.74 mph) regularly supplied for the DT-54.

The second modification, designated DT-57, is also made by the Stalingrad Tractor Works for the operation of "one-way" plows, as required in level irrigated areas, where "dead furrows" and ridges must be avoided. The modifications made to the basic tractor consist of:

1. Provide a pair of hydraulically operated 3-point hitches, one at either end. Suitable 3-bottom, deep tillage right- and left-hand integral plows are mounted which throw the dirt in one direction only regardless of direction of travel.

2. Modify transmission by addition of a forward and reverse "shuttle," so that five forward and five reverse travel speeds are provided.

3. Provide dual seating and dual controls, so that the operator may sit facing in either direction.

4. Increase number of track links from 41 to 42. Increase diameter of the large track idler such that the track contact on the ground is lengthened from 63 to 85.5 in., which is the same as on the DT-55.

The DT-57 tractor, equipped with the two plows

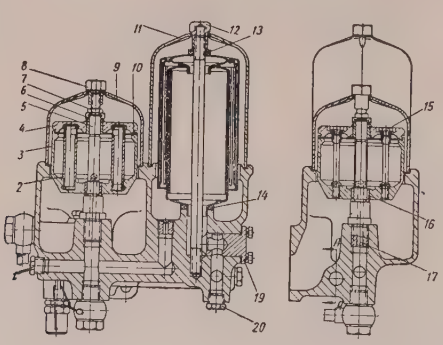


Fig. 27 — Reaction-driven engine oil centrifuge

is shown in Fig. 29. Should the design of this tractor fully meet field requirements, this particular model will probably be replaced by a new and improved tractor incorporating the same principals at the same time that the DT-54 is replaced by the DT-56, which is discussed later.

The use of opposite facing driver's seats involves a complicated control system, that is, gearshift, steering clutches, and the like, but these are evidently accepted (Fig. 30). However, the operator's vision is seriously impaired, especially with respect to plow mounted on the engine end of the tractor. Just how readily an operator adapts to these conditions and becomes qualified to do a satisfactory tillage job is questionable.

In 1953, there was introduced a modification of the basic DT-54 tractor equipped with a gas producer for operation on charcoal, designed Model GB-58. The compression ratio of the engine was reduced to 8.5/1 and spark ignition used. As compared with the DT-54 diesel tractor, the drawbar power was reduced by 10–12%, and the shipping weight was increased from 11,325 to 13,040 lb, of which only 594 lb is represented by the weight of the gas producer. With the more recent increase in petroleum production and better distribution facilities for diesel fuel, the need for operating on producer gas no longer exists.

"Stalinetz," Model S-100 — This large 100-hp crawler tractor is well-suited for pulling large ditching plows and drainage tools in areas where irrigation is practiced At the Northern Caucausus Machine Testing Station, one of these tractors had been provided with a heavy 3-point implement hitch, with hydraulic lift, and an 8-bottom semi-integral rigid frame plow, intended for use in the extensive level fields of that area. In the farm areas visited, few of these tractors were in use.

In general, the design of the S-100 is quite similar to the wartime "Caterpillar" tractor, many of which were shipped to the USSR. For this reason, no description or other data is needed, other than that given in the specifications.

New Developments

At present, plowing in most of the Soviet Union averages 8½–10 in. deep, except for sugar beets, where plowing depth ranges from 12½–13¾ in. and in the newly opened wheat areas of Siberia, where plowing depth is 6½–8 in. Subsoiling practices, in which the land is chiseled to a depth of 15–17 in. (without being turned over), are coming into wider use. Intensive research work is underway in many places in an effort to increase field working speeds, especially of tillage implements. The result of Soviet research work closely parallels our findings. From 5000 to 6000 agricultural engineers of all types are being graduated annually from the Soviet Institutes and Universities and the number of people engaged in the design and development of farm machinery is far greater than in this country.

Another factor contributing to the development of farm machinery in the Soviet Union is that there is no need that such a development be "profitable" to the tractor or implement plant engaging in this work. For instance, the acreage devoted to castor bean production is scheduled to be increased 10 times within the immediate future, a measure which requires the mechanization of castor bean harvesting. In a crash program, this task was assigned to institutes, design bureaus, and the engineers of combine plants and Machine Testing Stations, all of whom are required to give this project high priority. Under such conditions, rapid progress is inevitable.

Late in 1958, a similar crash program was undertaken under the sponsorship of GOSPLAN, the State Planning Authority, to convert the large drawn-type grain combines (of which there are over 200,000 units) into self-propelled units. A Stalin prize, in the amount of $20,000 was offered to the individuals or group whose design proposal was accepted. No information is available regarding the present status of this program.

Many interesting prototypes and pre-production tractors are under large scale testing or in pilot production. These developments are in the following areas:

1. Increase in power, with a higher power-to-

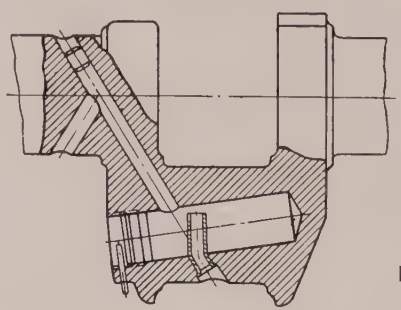


Fig. 28 — Crankpins with dirt-collecting cavaties

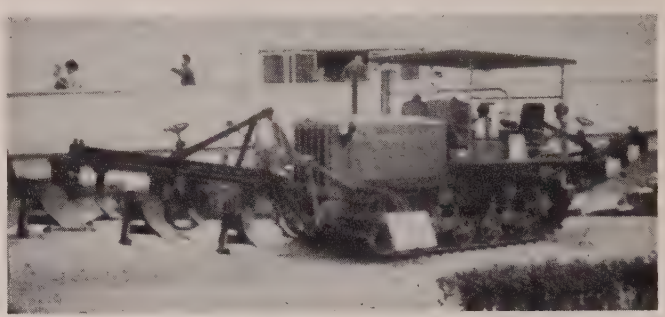


Fig. 29 — Model DT57 equipped with one-way plows

Fig. 30 — DT57 tractor, dual control and double seats

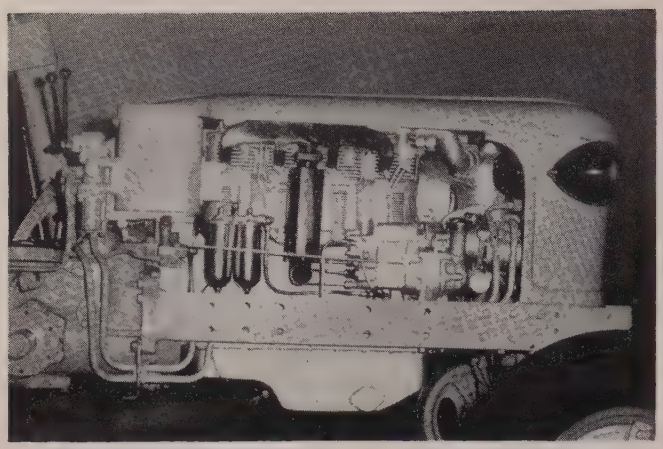


Fig. 31 — Four-cyl, 40-hp aircooled engine

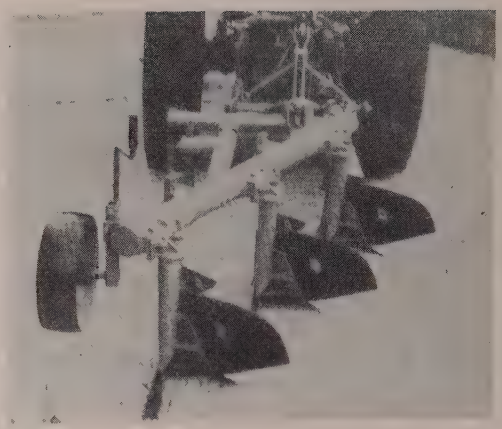


Fig. 32 — Three-bottom integral plow with welded tubular frame (rear view)

weight ratio.

2. Developments necessary for use with implements and machines suited for higher operating speeds.

3. Aircooled engines free from hazard of freezing.

4. Improved "implement carriers," otherwise termed "self-propelled chassis," which are intended for use with a wide variety of tillage and harvesting machines.

Increased Power — Developments now in progress indicates that the power of Soviet tractors can be expected to increase across the board. The following examples serves to show this trend:

1. A pilot quantity of 25 Model DSSH implement carriers has been produced, which are powered with 2-cyl aircooled 4-stroke diesel engines, 3.62 in. × 4.72 in., which develop 20 hp at 1750 rpm.

2. What appears to be a Model DT-28 chassis, powered with a 4-cyl aircooled diesel engine of the same series as above, has been exhibited. On the basis of 10 hp per cylinder at 1750 rpm, this will increase the power of this successor of the DT-28 from 32 to 40 hp. This engine is shown in Fig. 31.

3. A new 4-cyl engine designated D-50 has been under extensive development, as a replacement for the present D-40 engine used in the Byelaruss, Kirovitz KP-38, and other tractors. This engine has the same bore (4.33 in.), the stroke is shortened from 5.12 to 4.72 in. and the speed is increased to a contemplated 1800 rpm. This engine should conservatively develop 55 hp in place of the present 40 hp. Experimental tractors designated as E (for experimental) 50, using this engine have been built by the Dniepropetrosvsk Machine Works and tested at the Northern Caucausus Machine Testing Station. Very favorable performance has been reported.

4. Already announced is the DT-56 crawler tractor as a successor to DT-54. This tractor is powered with a lightweight 4-cyl diesel engine (4.52 × 5.11 in.) which develops 55 hp at 1500 rpm and 65 hp at 1700 rpm, with a claimed specific fuel consumption of 0.424 lb. The engine weight is 1210 lb, as compared with 2430 lb for the Model D-54 engine currently used in the DT-54 tractor; and the operating tractor weight is reduced from 11,890 to 10,615 lb. The engine is built by the "Sickle and Hammer Works" in Kharkov and is essentially the same as that used on the SK-3 Combine.

5. The Kharkov Tractor Works has developed a new engine, designated D-70, with the same bore (4.92 in.) as the present D-54 engine used in the DT-54 tractor, but with the stroke decreased from 5.98 to 5.51 in. This engine has a total weight of 1540 lb and develops 70 hp at 1500 rhm. A crawler tractor powered with this engine was extensively tested in 1958 at engine speeds up to 1800 rpm.

6. The Lipetz Tractor Works, whose product has previously been limited to crawler tractors, has displayed an "implement carrier" powered with the Model DV aircooled 4-cyl engine earlier developed by the State Scientific Tractor Research Institute. This engine employs a single blower, similar to that originally developed by Klöckner-Humboldt-Deutz in Germany. This engine has a bore and stroke of 4.13 in. and 5.11 in., respectively, weighs 1144 lb, its rated output is 50 hp at 1500 rpm, and the specific fuel consumption is 0.408–0.413 lb. Since the piston speed is a conservative 1290 fpm, a substantial increase in power is possible.

7. The Vladimir Tractor Works has also shown a conventional standard tread wheel tractor powered with this same 4-cyl aircooled engine. its size and weight comparable with the standard tread Model DT-24.

8. The Tchelyabinsk Tractor Works, whose largest production tractor is the Model S-100, has shown a new crawler tractor designated as the S-140. This is powered with a 6-cyl water-cooled diesel engine, 5.71 in. bore × 8.06 in. stroke (same as S-100), which develops 140 hp at 1100 rpm. This tractor uses low-cast manganese steel tracks and appears to be intended for agricultural and logging service.

9. "Rostelmash," the Rostov Combine Works, is developing a self-propelled implement carrier powered with the SMD-65 hp engine regularly used on their SK-3 combine. Several large integral implements including combine, corn picker, and hay baler have been experimentally built for use with this machine.

10. Minsk Tractor Works has undertaken the design of a self-propelled implement carrier powered with the Model D-40 engine used in the Byelaruss tractor.

Many prototype implements of advanced and clean cut design are in operation at Machine Testing Stations throughout the country. Figs. 32 and 33 show integral plows with a welded tubular frame. Although functionally deficient, the simple clean

Fig. 33 — Three-bottom integral plow with welded tubular frame (side view)

Fig. 34 — "Mechanized" sugar beet topping and cleaning after mechanical harvesting

frame design evidences original thinking and a new approach to tillage tool design.

In the Appendix are several illustrations showing the latest Soviet thinking in regard to farm machinery. Production items include the TDT-60 logging tractor, the DT-28 general-purpose wheel tractor, the DT-54 and DT-56 tracklayer.

Other advanced work in tractor design, including infinitely variable speed transmissions, is in progress at the State Scientific Tractor Research Institute, while similar work on implements is being handled by the All-Union Research Institute for Farm Machinery building. Both are directly responsible to GOSPLAN, both operate under the State budget. Many other institutes are engaged in similar activities.

Little additional information is available regarding the tractor research activities, but the Farm Machinery Institute has a total staff of 790 people, which includes 350 graduate engineers and 47 with advanced engineering degrees, of whom three are with doctorates. Of the engineers, 90 are women. In the year ending August, 1958, 70 prototype machines were built and tested in cooperation with implement factories throughout the country. The magnitude of this work is appalling, especially as compared with similar endeavors in this country. For instance, consider the Institute Strain Gage Laboratory, which is one of 15. This single laboratory alone has a staff of 20 people, 10 of whom are designers working on test equipment. This is twice the total number of people employed at the Soils Laboratory of the U. S. Department of Agriculture, at Auburn, Alabama. And the Auburn annual budget is $77,000 for personnel and $6000 for materials.

But in spite of brilliantly functioning Institutes and the finest in research facilities, the overall design and development of Soviet farm tractors and machinery is complicated, devious, and badly fouled up with red tape. There is a lack of continuing responsibility, as the following procedure clearly shows. All research is done under the direction of the Academy of Science of the USSR and the lesser academies of the individual republics. The Institutes develop principles and theory, and usually prepare working models, which are then field tested by Machine Testing Stations around the country. Following successful field test programs, projects are then turned over to specified Construction Bureaus for preparation of production designs. Some of these construction bureaus are located at the implement factories, but others are not. The individual constructing bureaus are responsible to the Council of National Economy, which in turn is responsible to the GOSPLAN of the Republic or area in which it is located. These, in turn, are responsible to the GOSPLAN of the Council of Ministers of the USSR. Before any new tractor or other farm machine can be placed in production, it must have had the approval of the Scientific Council of the Ministry of Agriculture.

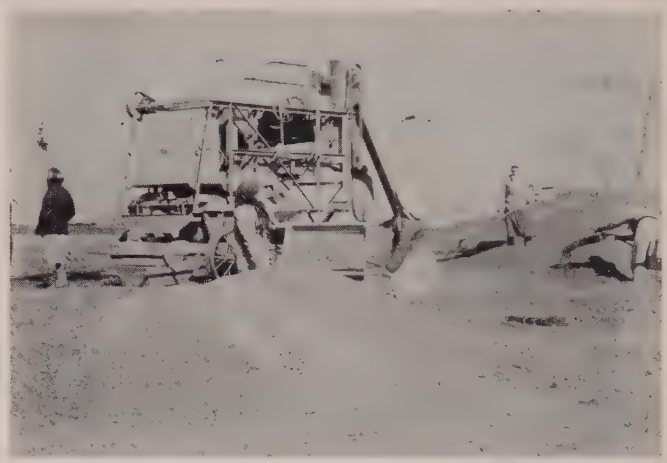
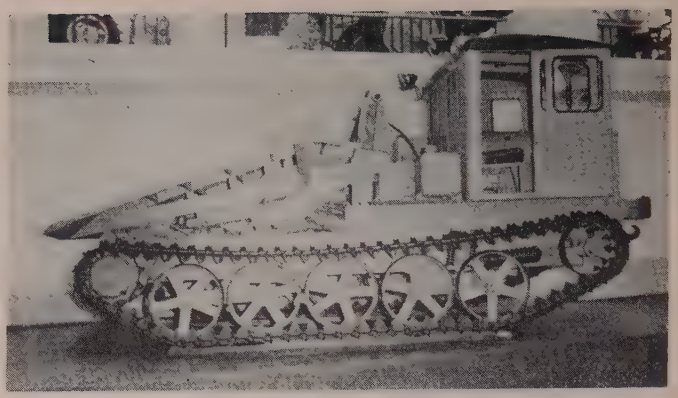


Fig. 35 — Grain cleaning on Soviet state farm (Soukhog)

At best, the mechanization of agriculture in the USSR is spotty. At one large Siberian State Farm (Sovkhoz), the harvesting of sugar beets was described as 100% mechanized. Tractor-drawn beet harvesters were used to top, lift, and pile beets and tops. A crew of three men was required: one tractor driver, one man on the beet harvester to guide it down the rows, and one "operator" who was busy keeping the harvester unclogged and who manually dumped the beet and beet top baskets when filled. In this country, one man operates both the tractor and the beet harvester and the beets are mechanically loaded into a truck driven along side the beet harvester. But on the Sugar Beet State Farm, it was necessary to re-top and clean the piled beets manually. Fig. 34 shows a crew of four women doing this work, with a labor expenditure of 25.4 (wo)man hr per acre. And this for an operation

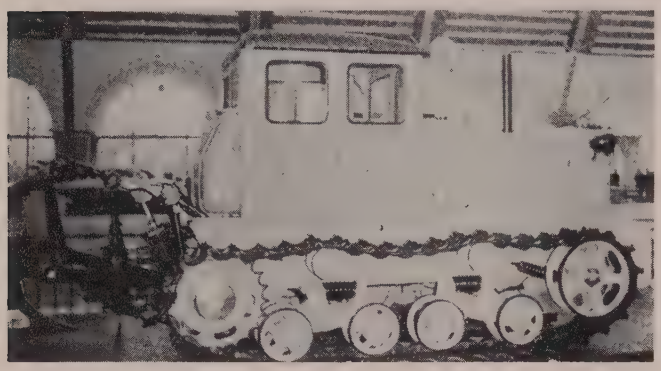
described as 100% mechanized! A similar "mechanized" operation is shown in Fig. 35. After combining, and before delivery to the elevator, grain is piled on a cement apron and manually handled into and out of the grain cleaner.

The degree and extent to which Soviet Agriculture is mechanized is truly difficult to define and appraise. But one simple comparison is enlightening. It is officially stated that in the Soviet Union, 56% of the population live on farms. In the United States, although no exact figures are available, the farm population is approximately 12% of the total. This small percentage of our people not only produce food and fibres for this country, but great quantities are exported and a surplus is accumulated which hangs ominously over our economy. On the other hand, free to draw upon the best thinking of the Western World, and with manpower and equipment for research and design and development of farm tractors and machinery far greater than in this or any other country, the rapid progress of Agricultural Sciences in the USSR can be readily understood. It may perhaps be best thought of as mobilization for the inevitable war of "peaceful competition."

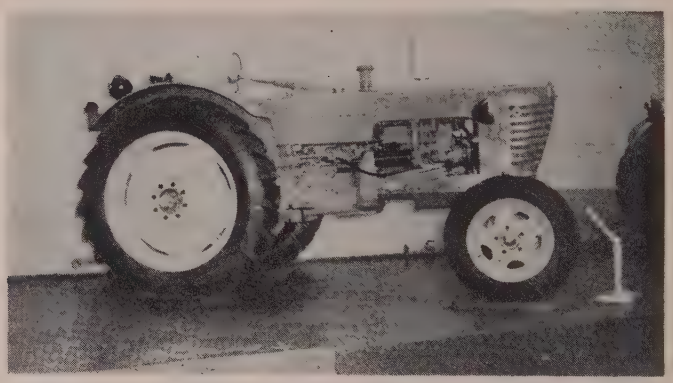


TDT logging tractor

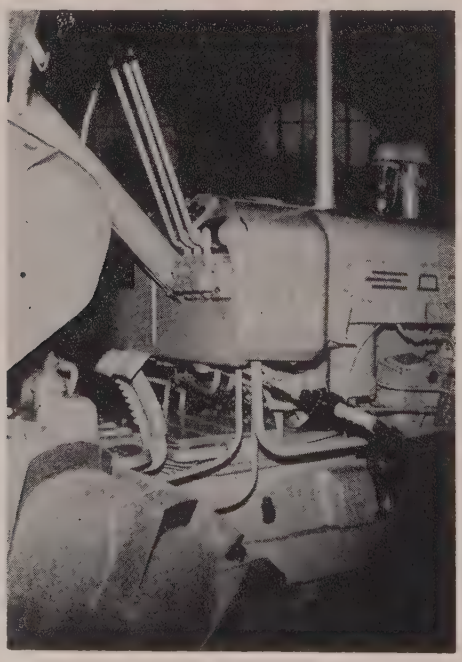
DT56 tracklaying tractor with 3-point hitch and integral plow



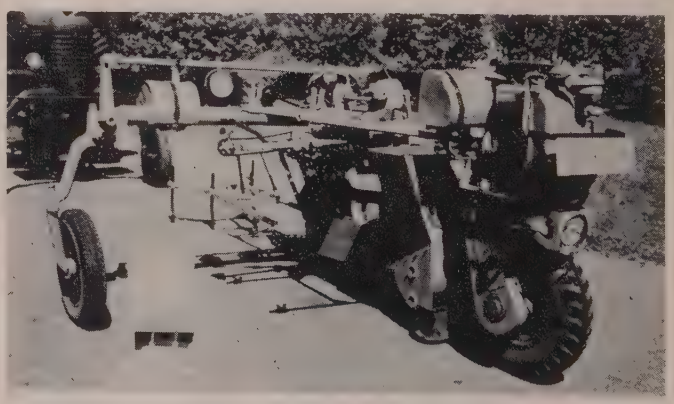
DT54 tracklaying tractor with 3-point implement hitch



Byelaruss MTZ-5 1959 model with 3-point hitch and remote hydraulic cylinder

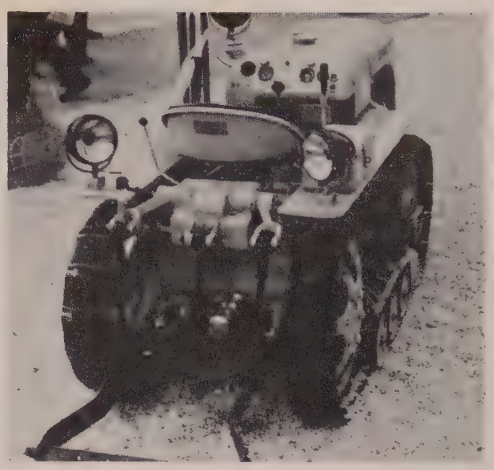


DT28 general-purpose tractor arrangement of three function hydraulic controls



Small front-wheel drive hillside tractor, prototype

Byelaruss MTZ-5 1959 model (rear view showing 3-point hitch with double-acting hydraulic lifting cylinder



Sixteen - hp Model DT-16G prototype tracklayer (Kharkov Tractor Works)

APPENDIX

Soviet Starting Engine Specifications

	PD 10(M)	P46		PD 10(M)	P46
Output, hp	10	17	Ignition	Magneto	Magneto
No. Cylinders	1	2	Lubrication	Oil Mixed with Fuel	Splash
Stroke	2	4			
Speed	3500	2600	Main Bearings		
Fuel	Gasoline	Gasoline	Number	2	2
Bore, in.	2.84	3.62	Type	Ball	Steel Backed Split Bushing
Stroke, in.	3.35	4.01	Crankpin Bearing	Roller	Steel Backed Split Bushing
Displacement, cu in.	21.1	83.1	Piston Pin Bushing	Bronze	Bronze
Compression Ratio	6.2	4.6	Starting Means	Rope	Hand Crank
Maximum Torque, lb-in.	186	615			

Soviet Tractor Engine Specifications

	D20	D28	D40	D54	KDM 100
Output, hp[a]					
Rated	18	28	40	54	90[b]
Maximum	20	32	45	—	100
No. Cylinders	1	2	4	4	4
Bore, in.	4.92	4.92	4.33	4.92	5.7
Stroke, in.	5.51	4.92	5.12	5.98	8.07
Displacement, cu. in.	104.1	187.5	274	456	828
Speed, rpm	1600–1800	1400	1500	1300	1050
Piston Speed, ft/min	1470–1650	1148	1279	1292	1411
Compression Ratio	15 & 5.6[c]	14.5 & 6.0[c]	17	16	15.5
Combustion System	Open	Swirl chamber	Swirl chamber	Swirl chamber	Precombustion chamber
Main Bearings					
Number	2	2	5	5	5
Type	Ball	Ball	Steel back	Steel back	Steel back
Material			Leaded bronze	Leaded bronze	Leaded bronze
Pistons					
Material			Aluminum alloy—all models		
No. Rings	4	5	5	5	4
Chrome Plated	1	1	1	1	1
Plain	2	3	3	3	2
Oil	1	1	1	1	1
Generator			12 volt — all models		
Starter	12v electric	12v electric or hand crank	12v electric or model PD starting engine	Model PD starting engine	Model P 46 starting engine
Injector					
Type	Pintle		Pintle injectors—all models		
Spray Angle, deg	15	15	15	15	—
Injection Pressure, psi	1775	1775	1775	1775	1705
Fuel Filter			Three-stage fuel filters—all models		
Oil Filter	[d]	[d]	[d]	[d]	[e]
Lubricating System			Pressure circulating—all models		
Rate, gpm	5.28	3.44	9.25	12.65	8.82
Oil Pressure, psi	25.6–36.4	25.6–31.3	28–42	24.1–35.4	7.0–38.4
Crankcase Capacity, gal	1.21	2.06	4.23	6.61	

cont. on p. 498

Soviet Tractor Engine Specifications, cont.

	D20	D28	D40	D54	KDM 100
Air Cleaner			Oil washdown with inertia precleaners—all models		
Bmep, psi					
Rated	82.8	84.4	77.1	72.1	82[b]
Maximum	84.8	96.5	86.8	—	91
Cooling Medium			All water-cooled		
Capacity, gal	1.95	4.0	6.6	15.9	16.9
Weight, lb	815	1100	1518	2795	4735
Lb/hp (rated)	45.2	39.4	37.9	51.8	47.4[b]
Specific Power					
Hp/cu in. Displacement					
Rated	0.173	0.150	0.164	0.119	0.109
Maximum	0.192	0.170	0.196	—	0.121[b]
Specific Fuel Consumption lb/hp hr	0.44	0.43	0.43	.48	0.44–0.46
Torque, lb in.					
Rated	708	1260	1680	2620	6000[b]
Maximum	779	1470	1990	3030	6230
Rise, %	10	14.2	18 6	15	4
Speed at Maximum Torque, rpm	1100–1300	900–1100	1100–1300	900–1000	700–750
Droop, %	19–31	21	13–27	23–31	28–33

[a] Output in terms of metric horsepower observed at engine.
[b] Continuous duty.
[c] Lower compression applies at time of starting.
[d] Three-stage oil filtration: wire wound element or cartridge, reaction centrifugal, and centrifugal chamber in crankpins.
[e] Three-stage oil filtration—twin cartridge with centrifugal chamber in crankpins.

Soviet Wheel Tractor Specifications

	DT-20	DT-24-1	DT-24-2	T-28	DT-24-3[a]	MTZ-5L MTZ-5M[b]	MTZ-7
First Production Year	1958	1955	1956	1959	1955	[c]	
Type	Universal	Tricycle -2 front wheel	General purpose	General purpose	High-clearance tricycle	General purpose	4-wheel drive
Engine Model	D20	D24		D28		D-40	D40
Engine Power at Crankshart							
Rated	18	24	24	28		40	40
Maximum	20			32		45	45
Weight							
Shipping	3210	5300	5485	4730	5480	6200	6835
Operating	3480	5765	5950		5940	6660	7315
Tractor Weight/Engine Power	178.5	221	228	148	228	155	170.5
Weight Distribution (operating)							
Front							
Weight	1280	1695	2035		1835		
Per Cent of Total	36.8	29.4	34.2		30.9		
Rear	2200	4070	3915		4105		
Wheelbase, in.	56-64.2-70-72.1[d]	87.9	88.9		87.9	93.5	
Tread, in.							
Front	43.3–59.1	8	49.2–68.9		49.2–78.75-94.5	47.3–70.8	52.5–57.3
Rear	43.3–55.1	47.2–70.8	47.2–70.8	47.2–70.8	70.8–78.75–91 6	47.3–70.8	51.2–70.8
Tires							
Front							
Number	2	2	2	2	1	2	2
Size	4.00–16	5.50–16	5.50–16	6.50–16	6.5016	605.–16	7.50–20
Rear	8.00–32	11–38	11–38	11–38	8.25–40 (14–24)	12–38	12–38
Turning Radius—Minimum ft	13.1	10.2	11.5		9.85	12.18	
Clearance under Rear Axle							
Quills, in.	18.1 or 20.25[e]	25.95		25.2	35.4	26.6	26.6
Clearance under Front Axle, in.	12.15		20.6			20.6	
Travel Speeds, mph							
Creeper							
1	0.53	0.36	0.36	0.298		0.85	0.85
2		0.78	0.78	0.497		1.05	1.05
3				1.43		1.34	1.34
4						2.19	2.19
5						3.00	3.00
6							
Forward							
1	3.14	2.92	2.92	2.27		3.93	3.93
2	4.06	3.36	3.36			4.84	4.84
3	5.12	3.98	3.98			6.06	6.06
4	9.78	4.55	4.55			10.05	10.05
5		11.75	11.75			13.9	13.9
6				15.5			
Reverse							
1	3.14	3.48	3.48	2.89		0.65	0.65
2	4.06	4.05	4.05	4.03		2.95	2.89
3	5.12						
4	9.78						

cont. on following page

	DT-20	DT-24-1	DT-24-2	T-28	DT-24-3[a]	MTZ-5L MTZ-5M[b]	MTZ-7
General Dimensions, in.							
Height							
Top of Exhaust		90.5	90.5		98.1	96.5	96.5
Top of Hood	48.5 or 56.5	60.7	60.7	61.8	68.9		
Width Over Rear Axles	51.4	76.3	76.3	76.3	83–91.5–108	74.2	74.2
Length		138.5	139.8	134.45	139.8	147.5	158.3
Belt pulley							
Speed, rpm	914	715	715	7.5	715	823	823
Diameter, in.	11.8	12.6	12.6	12.6	12.6	11.8	11.8
Width, in.	4.72	7.9	7.9	7.9	7.9	7.88	7.88
Belt Speed, fpm	2820	2360	2360	2360	2360	2540	2540
Number of Lights	3	4	4	4	4	4	4
Power Takeoff							
Type	Transmission driven	Direct engine driven	Direct engine driven	Direct engine driven	Direct engine driven	Direct engine driven	Direct engine driven
Speed, rpm	545	535	535	535	535[f]	520	
Speed of Ground Synchronized Power Takeoff, rpm							
Creeper		12.5 & 13.5					
1		90					
2		104					
3		123					
4		141					
5		365					
Tank & Reservoir Capacities							
Diesel Fuel, gal	12.15	18.5	18.5		18.5	26.5	
Starting Fuel, qt	0.8	1.7	1.7		1.7	0.79	
Engine Lubricating Oil, gal	1.19	2.06	2.06		2.06	4.23	
Cooling System, gal	1.96	3.96	3.96		3.96	6.61	
Transmission Case & Final Drive, gal		8.97	8.97		8.97	12.15	
Final Driver Gear Housing, qt					3.18		
Fuel Tank Capacity, hr at Full Load	13	11	11		11	13	13
Implement Hitch		3 point	3 point			3 point	
Drawbar	Swinging	Swinging				Swinging	
Hydraulic System[g]							
Pump	Gear			Gear		Gear	
Speed, rpm	1600			1400			
Output/min, gal	4.43			10.6		1500	
Rated Working Pressure, psi	1420			1350		1350	
Relief Valve Opening Press, psi	1775–1850			1700		1700	
Control Valve Positions	Neutral-Raise Lower-float			Neutral-raise Lower-float		Vertical-raise Lower-float	
Working Cylinder							
Diameter, in.	2.95			3.54		3.94	
Stroke, in.	4.33			7.87		7.87	

[a] Rear wheel tread adjusted by moving rear axle quills in and out by means of hydraulic cylinder.
[b] When equipped with electric starter, designated as MTZ-5M.
[c] Also delivered as Model MTZ-5 with starting engine and 5 forward speeds of 3.25, 3. 72, 4. 27, 4.91, and 8.62 mph.
[d] Wheelbase with drop housing toward 90 deg vertical back 45 deg or back 90 deg.
[e] Clearance under transmission case.
[f] Side pto available for seeders.
[g] Tractors of DT24 series and T28 optionally provided with up to three remote hydraulic cylinders for implement control.

Soviet Tracklayer Tractor Specifications

	KDP-35	T-38	KDP-35-2	DT-54	DT-55	DT-57	S-100	S-140
First Production Year	1950	1959.	1954	1949	1956	1957	1957	1958
Type	High clearance, general purpose	High clearance, general purpose	High clearance, cultivating	General purpose	Swamp & muskeg.	Two-directional tillage	General purpose	General purpose
Drawbar hp (rated)	24		24	36	36	36	62	
Engine Model	D38	D38	D38	D54	D54	D54[a]	KDM-100	6 KDM-50
Engine Power at Crankshaft, hp								
Rated	38	40	38	54	54	54	90	
Maximum	42	45	42				100	
Weight, lb								
Shipping	9140		9450	10710	12750	14520	95320	
Operating	9520		9900	11325	13420	14960	26100	33200
Tractor weight, lb/Engine hp	240		249	199	236	271	251	238
Wheelbase (centerline drive sprocket to centerline front idler), in.	67.2		82	102.5	105.5	105.5	93.25	
Tread of Tracks, in.	52.7	52.7	52.7	56.5	62	62	74	80.3
Ground Clearance, in.	25.2	25.2	31.5	10.25	10.6	11.1	15.4	

cont. on p. 500

Soviet Tracklayer Specifications, cont.

	KDP-35	T-38	KDP-35-2	DT-54	DT-55	DT-57	S-100	S-140
Travel Speeds, mph								
Creeper								
1	0.51		0.52	0.30	0.30	1.49–1.51[b]	None	
2	1.58		1.62	0.39	0.39			
3			0.45	0.45				
4			0.52	0.52				
5			0.99	0.99				
6			1.29	1.29				
7			1.50	1.50				
8			1.74	1.74				
Forward								
1	2.37	2.35	2.43	2.23	2.23	2.23	1.47	1.48
2	2.88	2.86	2.95	2.89	2.89	2.89	2.45	2.62
3	3.24	3.22	3.32	3.38	3.38	3.38	3.81	3.61
4	3.85	3.78	3.91	3.90	3.90	3.90	4.01	4.80
5	5.65	5.61	5.8	4.91	4.91	4.91	6.94	6.79
Reverse								
1	2.2	2.19	2.25	1.49	1.49	2.25	1.74	1.66
2						2.91	2.78	4.25
3						3.40	3.32	
4						3.93	4.68	
5						4.95		
Overall Dimensions, in.								
Height								
Top of Exhaust	108.1	103.5	111.3	90.6	90.6	93	120.3	110.5
Top of Hood	67.7		73.0	62	92	60.8	79.3	
Width	66.4	64.6	62.0	73.4	73.4	73.4	96.8	107.8
Length	122.2	143.2[c]	132.5	144.1	156.2	206.25	167.8	200.85
Implement Hitch	3 point	3 point	3 point	3 point		3 point each and attaches to hitch	3 point	
Drawbar	Rigid, adjustable	Rigid, adjustable	Rigid, adjustment	Rigid, adjustable	Rigid		Swinging	Swinging
Tracks								
Type	Semi-rigid	Semi-rigid	Semi-rigid	Articulated bogies	Articulated bogies	Articulated bogies	Semi-rigid	Articulated bogies
Track Rollers								
Number, Per Track	4		4	4	4	4	5	6
Mounting	Rigid		Rigid	Resilient-pivoting	Resilient-pivoting	Resilient-pivoting	Rigid	Resilient-pivoting
Idler Rollers								
Number, Per Track	1	1	1	2	2	2	2	3
No. of Teeth in Drive Sprocket	14	14	14	23	11	23	26	
Track Shoes								
Type	Cast steel with split bushings	Cast steel with split bushings	Cast steel with split bushings	Cast steel	Cast steel	Cast steel	Steel forgings — bolted	Cast steel — bolted
No. Per Track	34		37	41	40	42	41	
Width, in.	7.87 or 11.4	7.87 or 11.4	6.29	15.35	21	15.35	19.69	19.69
Pitch, in.	6.85	6.85	6.89	6.85	7.16	6.85	8.0	
Transverse Track Springs	Leaf	Leaf	Torsion bar[a]	None	None	None	Leaf	None
Length of Ground Contact, in.	30.75 aprox.	30.75	30.75	63	88.5	88.5	69 approx.	
Hydraulic System								
Belt Pulley								
Speed, rpm	690		None	764	764	Not supplied		
Diameter, in.	14.18			13.38	13.38			
Width, in.	7.87			9.83	9.83			
Belt Speed, fpm	2515			2675	2675			
No. of Lights	3		3	3	3	3	4	
Power Take-Off								
Type	Transmission[e] driven		Transmission[f] driven	Transmission driven	Transmission driven	Transmission driven	Transmission driven	
Speed, rpm	544 rear		527	547	547	547	1000	
Tank and Reservoir Capacities, gal								
Diesel Fuel	33.1		33.1	64.2	64.2	46.3	62.1[g]	
Starting Fuel	0.88		0.88	2.25	2.25	0.88	1.81	
Engine Lubricating Oil	4.5		4.5	6.61	6.61	6.61	7.15	
Cooling System	8.73		8.73	15.9	15.9	15.9	19.85	
Transmission and Final Drive	1.85		1.85	2.38	2.38	2.38	11.5	
Final Drive Gear Housing, each	1.32		1.32	0.45	0.45	0.45	5.81	
Fuel Tank Capacity								
Hours of Full Load	12		12	19	19	14	13.5	

[a] Basic D-54 engine with lubrication system modified to operate on slopes at 25 deg.
[b] Forward and reverse creeper speeds, respectively.
[c] Overall length with 3-point hitch.
[d] Experimental.
[e] Also provided with a transmission driven pto for mounted implement drive, operating at 5.37 rpm.
[f] Also provided with pto synchronized with travel speed turning at 45 rpm in first speed, or one revolution for each 4.64 ft of forward travel.
[g] 29-gal auxiliary diesel fuel tank available.

IH High-Speed Lightweight Diesels

A. F. Dewsberry,

C. P. Bozos,

and J. B. Reeves, Jr.

International Harvester Co.

THE NEW DIESEL ENGINES developed recently by International Harvester are adaptations of existing carbureted truck engines. The manufacturing economies possible by employing existing machine tools were the governing consideration in their design.

This paper discusses solutions found to the many problems presented by this approach to engine design: crankshaft, crankcase, connecting rods, pistons, lubricating system, and the like. Extensive interchangeability of parts was a result of the design concept.

The engines have been successfully adapted to farm tractors, crawler tractors, stationary motors, and various automotive applications.*

HIGH-SPEED operation and lightweight design have recently been coming to the aid of those time-proved diesel advantages of reliability and fuel economy. The resulting reductions in cost, weight, and package size have increased the versatility of the diesel engine, and have opened a multitude of new applications in all fields. Many new engines combining these features have made their appearance, leading a new expansion of the diesel industry.

The expansion has been very noticeable in the output range up to about 100 hp, covering most agricultural applications as well as a wide variety of automotive and industrial uses. The tremendous quantity of engines sold in this range brings the potential economies of high volume production within the grasp of the diesel builder. For example, 1958 diesel penetration of the farm tractor market reached 26% or 69,700 units, and March, 1959, tractor deliveries were 34% diesel.

Our company pioneered the tractor use of diesel power in 1933. The new family of diesel engines discussed here are the latest in the line — and are versatile power packages which are finding a wide variety of applications. (See Fig. 1.) The family consists of four 6-cyl engines covering the horsepower-speed range shown in Fig. 2, with displacements as follows:

236 cu in. — 3.688 bore × 3.688 stroke
252 cu in. — 3.812 bore × 3.688 stroke
282 cu in. — 3.688 bore × 4.390 stroke
301 cu in. — 3.812 bore × 4.390 stroke

Design

An apparent approach to the design of high-speed, lightweight, and low-cost diesels lay in the direction of adapting an existing truck engine to diesel combustion. Size limitations of our farm tractor chassis caused us to focus attention on the 6-cyl Black Diamond engine, our smallest truck engine made in 220, 240, and 264-cu-in. displacements. The 6-cyl in-line configuration would be completely balanced for primary and secondary forces and couples. Many production components and assemblies could be used. Existing machine tools had sufficient unused capacity to handle projected diesel requirements, a factor insuring a minimum capital expenditure *and* favorable unit cost. An alternate approach, of course, could have been the design of an entirely new diesel engine, and this received serious consideration. The manufacturing economies possible by employing existing machine tools, however, became the governing consideration.

Projection of future tractor horsepower requirements indicated a need for additional engine dis-

* Paper presented at SAE National Diesel Engine Meeting, Chicago, Oct. 27, 1959.

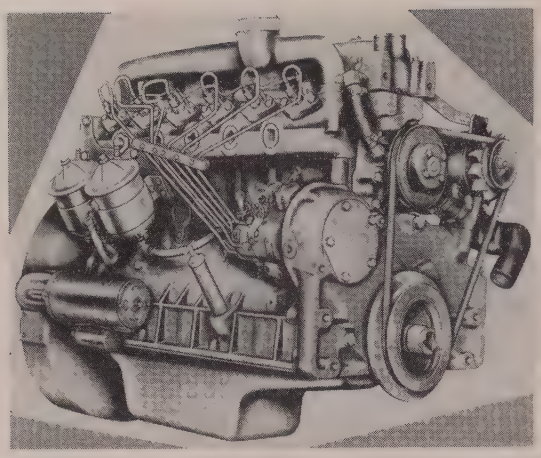


Fig. 1 — D-282 engine

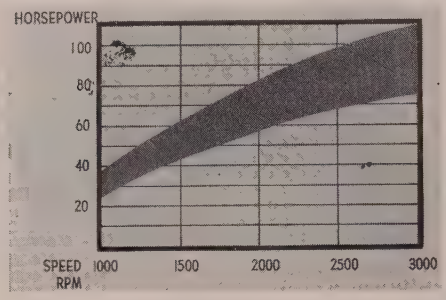


Fig. 2 — New IH diesel family:
3.688 × 3.688 – 236 cu in.
3.812 × 3.688 – 252 cu in.
3.688 × 4.390 – 282 cu in.
3.812 × 4.390 – 301 cu in.

placement. Detailed examination of crankcase clearances and related machine tools showed that the stroke could be increased from the existing 4.125 to 4.390 in., by lowering the crankshaft 0.218 with respect to the camshaft, without causing interference with the oil pump assembly. A sleeved bore of 3.688 in. was possible by revising water jacket cores, compared with the sleeveless Black Diamond bore of 3.562 (Fig. 3). This combination produced a displacement of 282 cu in. and was established as the basis of the new design.

The crankshaft lowering was accomplished with little effect upon most of the crankcase tooling. Additional stock was easily added to the bearing cap mounting face in the crankcase patterns. The cylinder boring equipment, and all drilling and tapping machines were located by dowel holes in the crankcase pan rail, thus no changes were required. The hydraulic broach for rough finishing the crank bore and cap mounting faces needed only the addition of new locating pads on its fixtures. The increased center distance between crankshaft and camshaft bores, of course, required purchase of new boring equipment. Finally, new bearing cap patterns were needed to prevent the caps from extending below the pan rail when assembled to the crankcase, and interfering with conveyors and fixtures.

Our Farm Tractor Division had found that the basic engine package could also be used in another, smaller, tractor. A stroke of 3.688 in combination with the 3.688 bore gave 236-cu-in. displacement and was selected to fit this power requirement. Concurrent with this diesel program and in very close cooperation, the Division was engaged in adapting the basic truck engine for lower speed operation, and with cylinder sleeves for more rugged service.

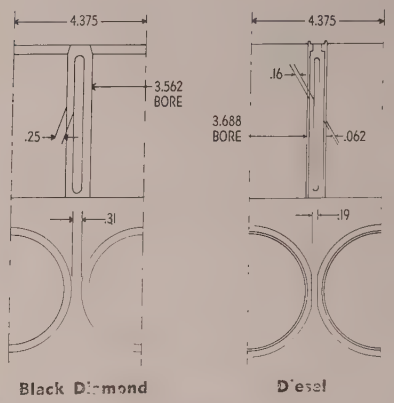


Fig. 3 — Crankcase section comparison

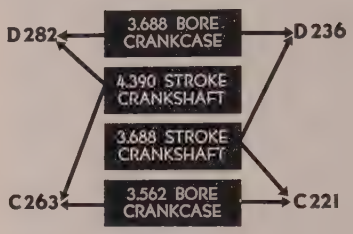


Fig. 4 — Engine family tree

As a result, these two diesel engines have close carbureted cousins related, as shown in Fig. 4. That is, two different crankcases and two different crankshafts were used to produce four separate engines.

A graphic summary of the use of common parts is made in Fig. 5, which shows a basic D-282 engine completely disassembled, with parts grouped as to common usage. Accessories and other items such as flywheels and flywheel housings are not included, since these are dependent upon particular applications and can fit any engine if desired. The D-252 and D-301 engines are not included in this summary, but will be described later.

Parts common to all engines in the family are grouped in the upper right of Fig. 5, and include bearings, camshaft, tappets, oil pump, water pump, oil pan, oil filter components, connecting-rod caps, and others. The crankshaft gear also belongs in this group, although shown assembled to the crankshaft. The left side shows parts common to the D-236 and D-282 engines and includes the majority of items. The cylinder heads for these two engines do differ in the diameter of discharge hole leading from the precombustion chambers, but the same castings and tooling are used. The lube oil filter base uses the same foundry equipment and tooling as the base for the carbureted engines, but is also machined for the oil cooler. Injection pumps for different applications will have some internal differences.

The crankshaft is common with the C-263 carbureted engine, leaving only pistons and precombustion chambers as parts peculiar to the D-282. Since the same connection rod and crankcase are used for both strokes, the D-282 and D-236 have different pistons. Both pistons, however, use the same rings, full-floating piston pin, and diametral dimensions and are machined on the same equipment. Even the precups have the same external dimensions as those for the D-236 and are machined on the same tooling.

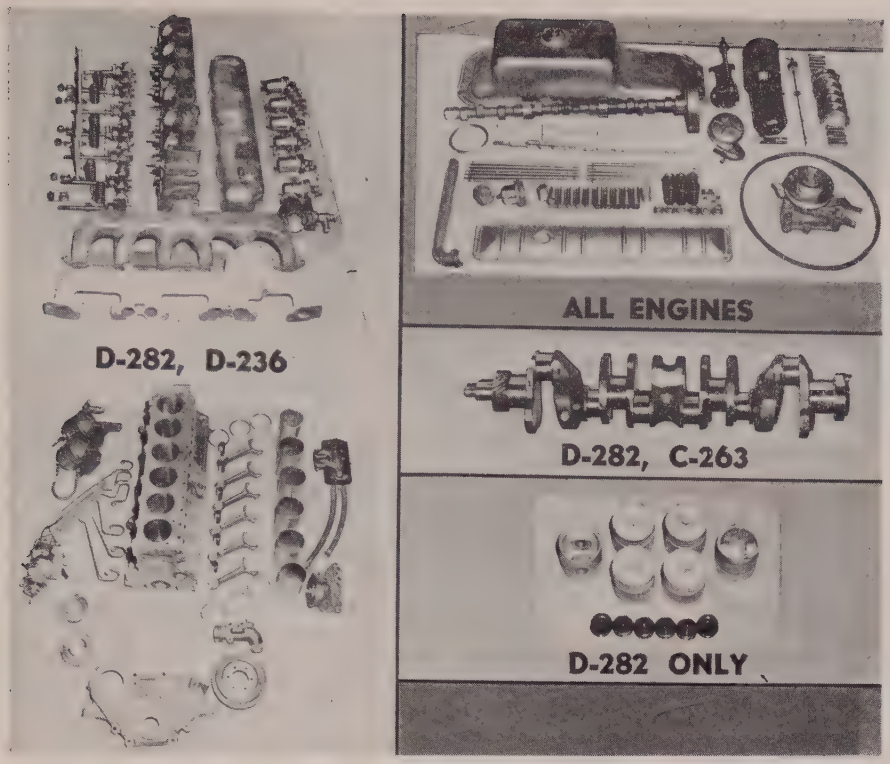


Fig. 5 — D-282 engine disassembled, showing common parts

This extensive use of interchangeable parts is a result of the basic design concept of this family, and enables our dealers and distributors to provide complete service with a minimum of parts inventory. Furthermore, this permits higher production rates and lower costs for these components.

Crankshaft and Bearings

The crankshaft was a key area in this program and should be considered in more detail. Use of existing machine tools dictated the cylinder spacing, the four main bearing configuration, and the main and rod bearing diameters. The bending stress level in the existing shaft, however, was excessive when projected for the increased bore and the higher cylinder pressures of a diesel. The decrease in effective section modulus due to the increased stroke also contributed to this condition. It was necessary, therefore, to increase cheek thickness to provide sufficient section modulus. Fig. 6 gives a comparison of the original 4.125-stroke gasoline crankshaft and the 4.390-stroke diesel shaft.

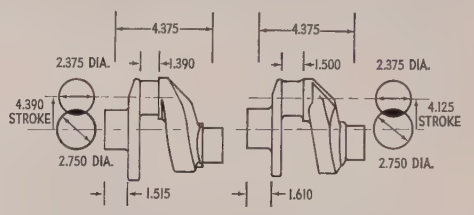


Fig. 6 — Crankshaft comparison (all fillets 0.170 R)

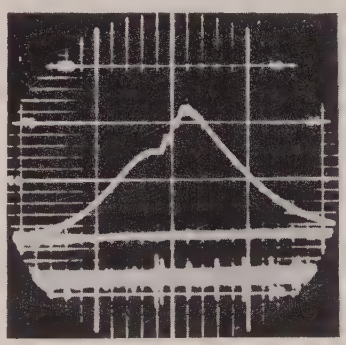


Fig. 7 — Pressure-time trace (110 bmep at 1800 rpm, peak 960 psi, vertical scale 85.7 psi/square)

Since the actual criterion of crankshaft design was the fillet bending stress, applied load was as important a consideration as section modulus. This dictated our selection of the precombustion chamber system, with a peak pressure of 960 psi (as shown in the pressure-time trace of Fig. 7), rather than a direct injection system with its higher pressures.

In order to use the same crankcase, bearings and other parts, the 3.688-stroke crankshaft has the same cheek thickness as the 4.390 and, consequently, operates at a much lower stress level. Both crankshafts are forged from C-1046 steel with a finished Brinell hardness of 249 to 298. Bearing journals are not induction hardened in keeping with the manufacturing procedure for the gasoline shafts, but are polished to assure long bearing life. The crankshafts are rifle drilled to provide pressure lubrication to main and rod bearings. Although new forging dies were necessary for both crankshafts, existing production machine tools were employed with changes only in setup and in perishable tools. The rifle drilling in the 4.390-stroke shaft was designed around the drilling machine used for the existing 4.125-stroke shaft. Front and rear ends of the diesel crankshafts are identical with the original gasoline shafts. The increased crankshaft cheek thickness also produced an anomaly in bearing areas. Since cylinder spacing was fixed, the cheek increase

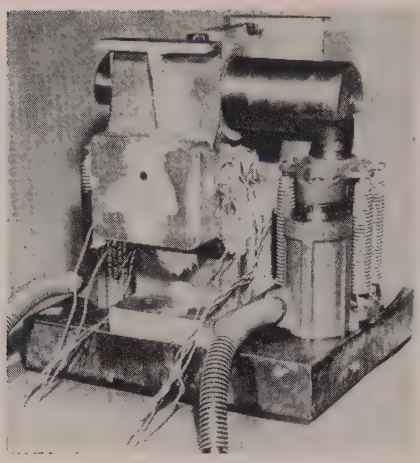


Fig. 8 — Crankcase and bearing cap test fixture

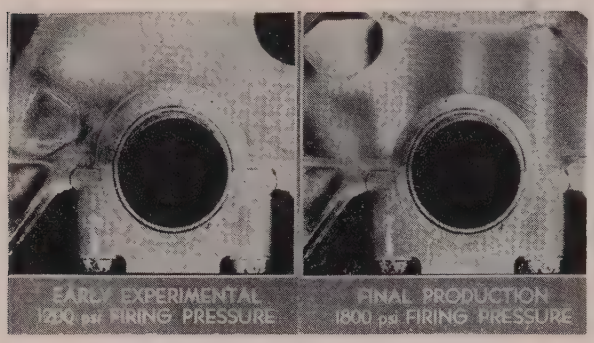


Fig. 9 — Crankcase bearing web modifications

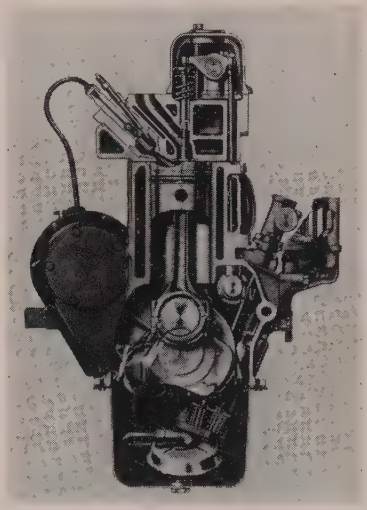


Fig. 10 — Typical engine cross-section

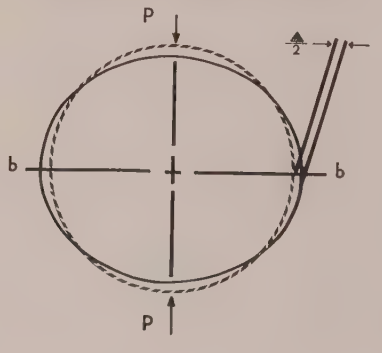


Fig. 11 — Piston pin deflection

was necessarily at the expense of bearing length (Fig. 6). This resulted in a diesel engine with higher applied load than the original gasoline engine and with less bearing area to carry it. A comparison of bearing areas and loads is made in Table 1. The trouble-free service records of the thousands of engines in service have proved the ability of the heavy-duty trimetal bearings (steel back, copper lead alloy center, babbit overlay) to carry these loads.

Crankshaft

Certain features of the crankcase have already been mentioned. The same pattern and cylinder barrel cores are used as for the carbureted engines, with different water jacket cores. These water jacket cores provide for the larger bore and also for the different location of some cylinder head capscrews necessary to install the precombustion chambers. Alloy gray iron is used for the crankcase, while the four main-bearing caps are made from high-strength nodular iron. The first experimental engines were successfully tested with gray iron bearing caps, but the subsequent use of engine speeds over 1800 rpm caused bearing cap failures which were eliminated by use of nodular iron. Capscrews are used to mount the bearing caps, and also the cylinder head, to the case.

The bearing cap support web was the subject of considerable stress analysis to determine its ability to sustain turbocharged combustion pressures. The crankcases used in the early experimental program were found capable of carrying 1200-psi combustion pressure and were acceptable for naturally aspirated operation. A static test setup was constructed (Fig. 8) using a full-scale cast-iron model of the bearing web, loaded with hydraulic rams to simulate firing pressure. Strain gages were located at critical spots revealed by a brittle lacquer, or Stresscoat, test. Use of this model, made from a simple wood pattern, enabled changes to be evaluated rapidly and inexpensively without disturbing the production crankcase pattern equipment. A comparison of the original web of 1200-psi capacity, and the final design capable of carrying 1800-psi combustion pressure is made in Fig. 9. These changes were made prior to the start of turbocharged engine production, and were also made in the related gasoline crankcases to maintain common foundry equipment.

The cylinder sleeves used with the crankcase are centrifugally cast alloy gray iron with a wall thickness of only 0.062 in. and are a hand-push fit in the crankcase bores. The sleeves have an as-cast Brinell hardness of 217–277 after machining, and no heat-treatment is used.

Piston and Connecting Rod

Piston and connecting-rod assembly is shown in Fig. 10. Connecting-rod design is similar to that of the Black Diamond engine, except that the I-beam section was increased to accommodate higher loads and the crank end made narrower due to the increased width crank cheeks. In addition, the connecting-rod length was increased 0.250 to 7.250 in. due to the relocated crankshaft center. The diesels also employ a larger diameter piston pin than the carbureted engines, for reasons to be discussed. These modifications required a new forging die, but the production machine tools are used with

Table 1 — Bearing Load Comparison

Bearing	Highest Peak Pressure, psi		Highest Mean Pressure, psi	
	Diesel (3.688 bore)	Gas	Diesel (3.688 bore)	Gas
Connecting Rod	5030	2720	990	515
Intermediate Main	3110	1470	800	380

only fixture and setup changes. Two self-locking capscrews are used to secure the cap to the connecting rod. Piston pin lubrication is provided through a hole drilled in the top of the connecting-rod pin boss.

The novel three ring combination used on the aluminum alloy piston was the result of laboratory tests of many combinations. Our objective was the use of MIL-L-2104-A lubricating oil for farm tractor applications. The low additive level of this oil caused top ring sticking and high blowby with the initial plain and taper faced rings. Pistons using four rings were tried, as well as combinations of different ring types. Top ring sticking was finally eliminated by use of a half-keystone section. The ferrox-coated intermediate ring is taper faced and the single oil ring uses a coiled spring expander with 250-psi radial load. The top and oil control rings are chrome plated for long life.

The piston pin was the cause of a detailed study, results of which will be summarized. Since the piston pin is one of the reciprocating elements, it is always made hollow to reduce inertia loads. Consideration of the pin as a beam will show that this affects its section modulus very little. However, its deflection as a hoop or thin-walled cylinder often becomes significant. Assuming a thin walled cylinder loaded as in Fig. 11, the increase Δ of the horizontal diameter may be found from theories of elasticity to be:

$$\Delta = 0.137 \frac{PR^3}{EI} \qquad (1)$$

Substituting in this expression as follows:

R = Mean pin radius = $\dfrac{D+d}{4}$

E = Modulus of elasticity = 30×10^6 psi

P = Applied combustion load

I = Section moment of inertia = $\dfrac{bh^3}{12} = \dfrac{L}{96}(D-d)^3$

where:

L = Pin length
D = Pin OD
d = Pin ID

results in the following expression:

$$\Delta = 68.5 \times 10^{-10} \frac{(P)}{(L)} \left[\frac{(D+d)}{(D-d)} \right]^3$$

This resembles very closely an equation given by Reiners and Schmidt,[1] their constant being 57×10^{-10}. No derivation was indicated, however. The relation just derived agrees very closely with actual measurement of deflection of piston pins between flat plates, as shown in Fig. 12. This curve is a plot of the equation in the form

$$\frac{\Delta L}{P} = 68.5 \times 10^{-10} \left[\frac{(D+d)}{(D-d)} \right]^3 \text{ versus } d/D$$

in order to show different values of P/L. The actual test data check closely with the curve for values of d/D greater than 0.50 and diverge as d/D decreases. This is to be expected, since decreasing values of d/D represent greater deviation from the assumed, thin-walled cylinder.

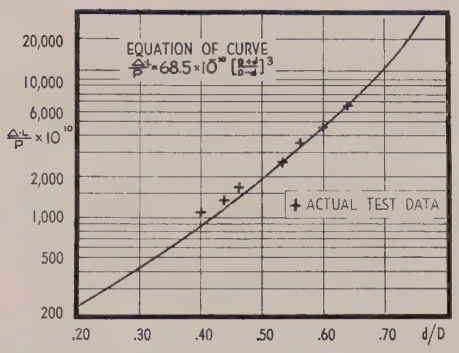


Fig. 12 — Piston pin deflection

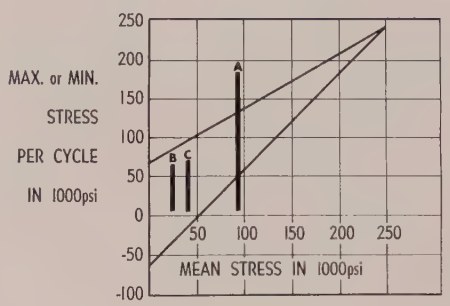


Fig. 13 — Modified Goodman diagram showing piston pin stress

Bending stress in the pin was next considered, using the value of the maximum moment for this type of loading to find stress, $S = Mc/I$. Substituting for R and I as before and with $c = (D-d)/4$:

$$S = 1.908 \frac{P}{L} \frac{(D+d)}{(D-d)^2} \qquad (2)$$

This value represents the bending stress on the inside and outside surfaces of the pin directly in line with the load.

A modified Goodman diagram for the piston pin material used is shown in Fig. 13. Line A indicates the stress condition (based on the preceding analysis) for the Black Diamond gasoline pin under diesel loading and predicts failure by splitting into barrel staves. The correctness of this prediction was strikingly verified by the inadvertent use of these pins in an early experimental engine. The original diesel pin had the same length and OD as the gasoline pin, but with a smaller ID. Line B indicates the condition for this pin, which was used throughout the test program and early production. The use of turbocharging caused a re-evaluation of this component due to the expected higher cylinder pressure. The decision was made, therefore, to increase the pin OD and to stiffen it by a thicker wall to Condition C. This change also insured sufficiently low bearing pressures in the piston bosses. To maintain interchangeable parts, this change in piston pin, rod, and piston has been made for all diesel versions of the engine family. Connecting-

[1] "Experimental High-Speed Cummins Diesel Engines, 4000 rpm, 345 hp, 401 cu in.," by N. M. Reiners and R. C. Schmidt. Paper presented at SAE Summer Meeting, French Lick, Ind., June, 1951.

rod caps and bolts remain common with the carbureted versions.

Cylinder Head

The cylinder head is a one-piece alloy gray iron casting (Fig. 10). For better volumetric efficiency, each cylinder has its own intake and exhaust ports with manifolds on opposite sides of the engine. Intake and exhaust port contours were determined with the aid of airflow measurements made on plastic models, prior to construction of the experimental cylinder head pattern. A baffle is cast horizontally across the head to direct coolant flow through the valve seat areas and to improve rigidity of the head. In addition, vertical walls are used to insure maximum coolant flow through each intake—exhaust bridge, rather than between cylinders (Fig. 14). All exhaust port cores are identical, while intake cores are right and left hand. The water jacket core is made in two pieces with supporting chaplets in the baffle. The head is cast bottom deck down in green sand molds made by pneumatic pressure in a diaphragm molder. Due to the extensive differences between the diesel and gasoline cylinder heads, an entirely new machining line was installed for the diesel head. The intake valve is made from Silchrome XB and is 1.652 in. in diameter. Rubber boots are used on the intake valve stems to prevent oil loss through the guide. The exhaust valve is made from 21-4N with welded stem tip and is 1.300 in. in diameter. Positive valve rotators are used with all valves. The stamped rocker arms are shaft mounted; the valve cover is aluminum die casting.

The cylinder head gasket was the source of many problems in the experimental program. The gasket has a steel core with prongs securing the treated asbestos outer layers. Steel overlaps are used around the combustion chamber, with copper grommets at water and oil holes. One group of water holes required use of rubber inserts.

Camshaft

The camshaft is the same as used in the gasoline engine and is a steel forging with induction hardened lobes. Valve timing is shown in Fig. 15, and is the same for all applications and speeds including turbocharged. The lube oil pump is driven by a right angle drive in the center of the camshaft with a tachometer drive on the outer end of the oil pump shaft.

Lubricating System

The lubricating oil pump is driven by a worm gear on the camshaft. The original Black Diamond gear-type pump is used without change. Oil is delivered through a flat plate cooler and a full-flow filter to the main oil gallery for distribution to bearings, gears and valve mechanism (Figs. 16 and 17). The filter retains particles larger than 25 microns. Oil pressure regulation takes place on the clean side of the filter, returning excess oil to the stamped oil pan. This arrangement prevents the regulating valve from being held open by dirt particles, permits filtration of more oil, and maintains a more constant gallery pressure. The cooler and filter are each equipped with bypass valves as safety devices to permit flow to the bearings if the cooler passages or filter element become plugged. The oil pump inlet screen, or floto, was revised slightly to permit operation of the engine at angles up to 45 deg.

Use of a flat-plate oil cooler in preference to the more conventional tube bundle was dictated by two considerations: (1) cost — since this cooler is a simple adaptation of an automotive transmission cooler, we were able to obtain the benefits of high production tooling; and (2) space limitation imposed by the steering linkages of our Farmall tractors. This full-flow system was designed for these engines as a result of very successful experience with other diesels in our line. Comparison of bearings from early experimental engines of this family run with bypass filter only, and with this full-flow system, strongly backed up this choice. Actual measurements showed main bearing shell wear of 0.0015 after 189 hr at full load with the bypass filter, and 0.0002 after 5930 hr with the full-flow filter. It was an interesting design problem to provide full-flow lubrication and cooling on a crankcase originally designed for bypass only, without requiring new machine tools. The carbureted tractor engines related to this family use the same full-flow filter without the cooler.

Cooling System

The original Black Diamond centrifugal water pump is used, modified only by connections for the

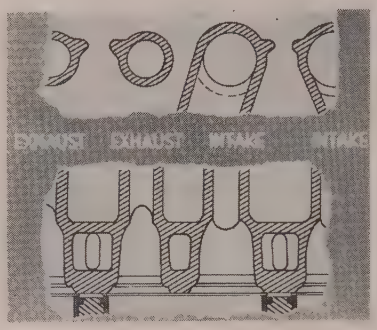


Fig. 14 — Valve bridge area

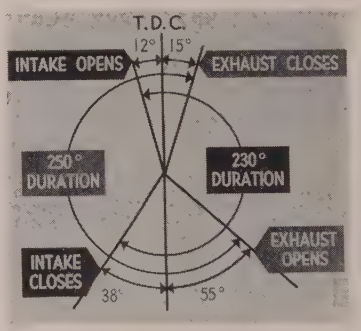


Fig. 15 — Valve timing diagram

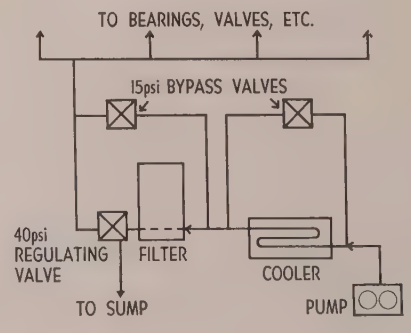


Fig. 16 — Lube system schematic

oil cooler and the thermostat bypass. Water is delivered directly into the crankcase, circulating around the cylinder barrels. Eighty per cent of the water comes into the head beneath the baffle, insuring maximum flow through the valve bridge area. Flow out to the radiator is through the thermostat by way of a front outlet, cast integral with the head.

Injection System

The injection system is the result of an intensive development program. As previously noted, a precombustion-chamber system was selected to insure minimum firing pressures. The precombustion chambers are considerably smaller in volume than previous IH diesel engines and, therefore, required development of a new injection nozzle. Initial engine tests with various nozzles determined the basic spray characteristics. Numerous nozzles were then bench tested to categorize the spray. From these tests certain nozzle and pump combinations were selected for further engine tests in which various configurations of precombustion chambers were tried. Over 200 variations of nozzles and pumps were either bench or engine tested. Optimum performance was obtained with these components:

1. Injection pump — plunger size matched to engine displacement, with delivery valve and automatic timing advance.
2. Injection nozzle — outwardly opening pintle type, 1525–1625 psi opening pressure.
3. Fuel line — 0.093 in. ID × 27.3 in. length.

The lightweight injection pump is flange mounted on the engine front plate. The pump drive presented an unusual problem, since space limitations would not permit a single gear twice the diameter of the crankshaft gear, as is customary with a half-engine speed pump. The fact that an idler gear was necessary between the pump and crankshaft gears enabled the solution. A double gear is used as the idler, with crankshaft gear driving the larger idler and the pump driving from the smaller (Fig. 18). Timing adjustments are made by turning the slotted gear about its pilot on the pump hub. The idler gear hub and shaft are used on the carbureted tractor engines with a single gear as a mechanical governor drive.

Two basic variations in automatic timing advance are used — "load advance" for tractor-type operation and "speed advance" for automotive operation. With either of these units, the cam is in the retard position during cranking and, therefore, greatly assists cold starting.

1. Load Advance — In the high swirl precombustion design used in these engines, it is necessary to maintain a constant start of injection to prevent misfiring during warmup, particularly at lower ambient temperatures and at speeds above normal idle. Since the basic injection pump has a constant end of injection, an automatic unit was developed to advance the cam as loads are reduced, thereby prevent misfiring. For these applications, the cam advances 8 engine deg between full load and no load at any engine speed (Fig. 19).

2. Speed Advance — On applications where rated speed is above 2400 rpm, there is a major advantage in having start of injection advance with speed. With this advance unit, the cam starts advancing at 1800 rpm and reaches full advance at 2900 rpm for a 3000 rpm rating. A total of 6 engine deg advance is used. The advance curve shifts slightly to lower speeds at no-load conditions, thereby maintaining satisfactory no-load operation (Fig. 20).

Two fuel filters capable of retaining 5 micron and larger particles are mounted to the cylinder head and are connected in series on the inlet side of the

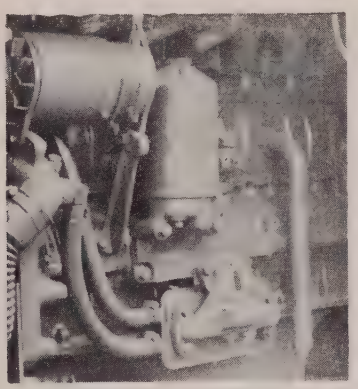


Fig. 17 — Full-flow oil filter and cooler

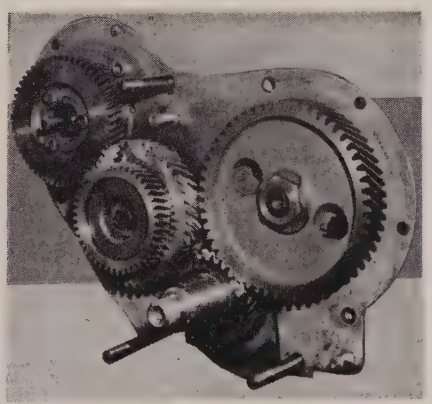


Fig. 18 — Engine gear train

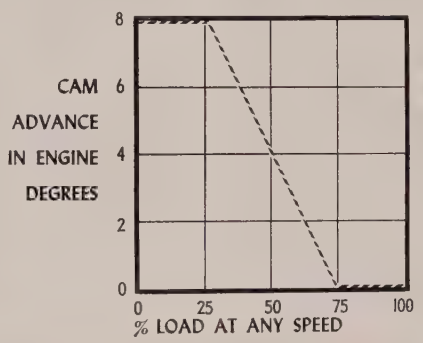


Fig. 19 — Load advance

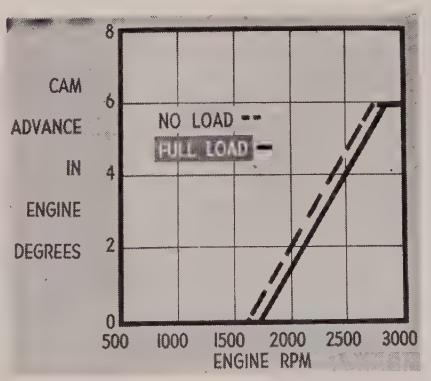


Fig. 20 — Speed advance

primary pump located on the end of the injection pump.

Starting

These engines are departures from previous IH diesel engines in that an integral gasoline starting system is not used. Instead, 12-volt glow plugs are located in each precombustion chamber. These are energized before cranking for 30–90 sec, depending upon ambient temperature. The glow plugs, together with an 18/1 compression ratio, have proved to be a fully satisfactory replacement for the gasoline system and at a considerable reduction in cost and complexity. Carburetor, distributor, spark plugs, starting valves, and the like are not needed on these new diesels.

This starting system had been thoroughly tested at our proving grounds through two winters prior to the start of production, as well as in our own cold-room facilities. In addition, thousands of engines in customer use successfully operated through the recent severe winter of 1958–1959. Typical starting data for a crawler tractor installation with 12-volt cranking system are as follows:

1. – 9 F — After 90-sec glow plug preheat, engine started in 3.5 sec. Engine serviced with 9 qt 10W oil.
2. – 26 F — After 90-sec glow plug preheat, engine started in 17 sec. Engine serviced with seven quarts 10W oil plus 2 qt kerosene.

During these tests, the entire tractor, batteries and fuel system were "soaked" at the temperature noted for 64 hr. As mentioned before, the injection pump cam is in a fully retard position during cranking, thus greatly assisting cold starting. In fact, at – 5 F, satisfactory starting is accomplished with 90 rpm cranking speed with pump cam in this retard position. If the cam is advanced 4 engine deg, cranking speed must be above 120 rpm.

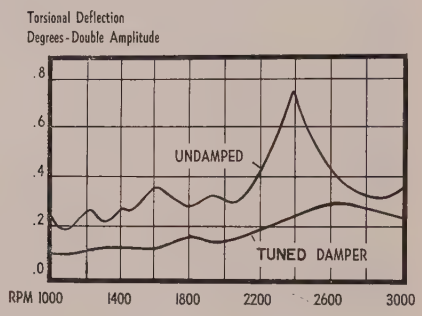


Fig. 21 — Torsional vibrational characteristics of D-282 at full load

Other Features

A torsional vibration damper incorporated in the front pulley is employed to limit crankshaft torsional amplitudes. Due to space restrictions, it was necessary to place the pulley groove itself in the inertia member. This novel arrangement has proved quite satisfactory, with the same inertia and elastic members used on all naturally aspirated engines regardless of speed. Torsional vibration amplitudes measured at the front of the crankshaft are shown in Fig. 21. The undamped curve shows a very high sixth order disturbance at 2400 rpm, while the damped curve shows the final result of tests of many types of vibration dampers.

Vacuum pump, air compressor, hydraulic pump, and alternate position generator are available as attachments. Provision has been made for a variety of cooling fan locations. A front power take-off connection capable of transmitting 30 hp is part of

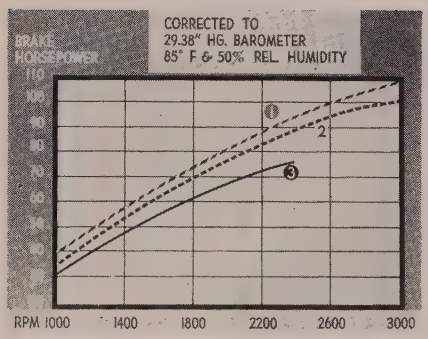


Fig. 22 — D-282 engine brake horsepower

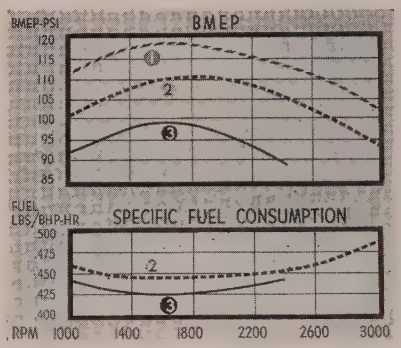


Fig. 23 — D-282 engine bmep and fuel consumption

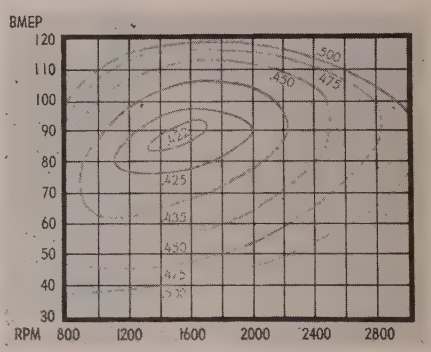


Fig. 24 — D-282 specific fuel consumption map, lb/bhp-hr

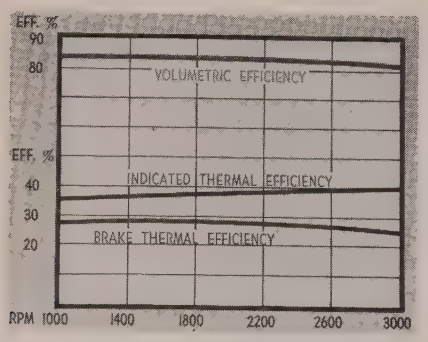


Fig. 25 — D-282 engine, volumetric and thermal efficiency (fuel 19,600 Btu/lb)

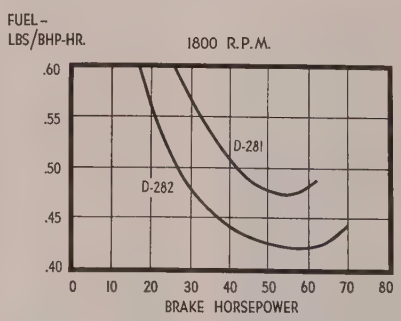


Fig. 26 — Comparison of 4-cyl D-281 engine and 6-cyl D-282 engine

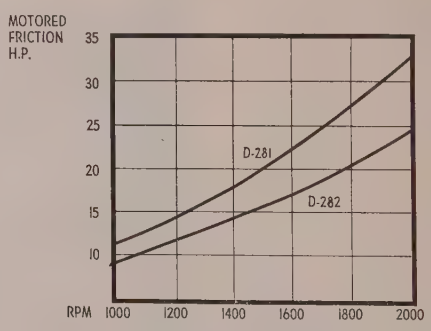


Fig. 27 — Comparison of 4-cyl D-281 engine and 6-cyl D-282 engine, friction hp

the front pulley.

The advantages of using existing machine tools have already been mentioned. One of the unusual problems this created was the effect upon diesel compression ratio of normal carbureted engine machining tolerances. The effect of connecting-rod length and crankcase height tolerances upon a 7.5/1 compression ratio is small, since most of the clearance volume is not dependent upon the piston-to-head dimension. These same tolerances on the 18/1 compression ratio diesel produced a wide variation, since the space between piston and head had become a major portion of the clearance volume. The effect of this compression ratio variation on the diesel engine was considerable, particularly in starting and idling characteristics. It was possible to reduce the connecting-rod length tolerance by 80% without additional cost, since the existing machine had this capability. The crankcase tooling could not achieve the desired tolerance, however, requiring purchase of a grinder and an additional machine operation for diesel crankcases.

Engine weight with fan, generator, and cranking motor ranges from 860 to 950 lb, depending upon end product application. Engine dimensions are 39.47 in. long (less fan), 21.28 in. wide, and 29.21 in. high.

Performance Data

The horsepower and speed ranges available from this engine family were shown in Fig. 2. For ease of presentation, specific performance data will be discussed for only one engine, the D-282. This engine is used in the greatest number of different applications and is the only turbocharged model to date. All naturally aspirated performance data are corrected to the latest standards recommended by the SAE Diesel Engine Test Code Committee of 29.38 in. of Hg, 85 F, and 50% relative humidity. These conditions are very close to the yearly average conditions in our Melrose Park, Ill., laboratory; consequently actual correction factors used in the data presented range from 0.984 to 1.009. Turbocharged performance has not been corrected.

Brake horsepower versus speed is shown in Fig. 22, with curves 1, 2, and 3 identified as follows:

Curve 1: Maximum horsepower — that which engine can develop for at least 5 min without drop in speed.

Curve 2: Intermittent horsepower — usable for variable load applications such as tractors, motor graders, air compressors.

Curve 3: Continuous horsepower — for steady day and night loads over long periods, such as irrigation pumping.

Specific fuel consumption and bmep corresponding to these are shown in Fig. 23. Maximum torque for the D-282 is 220 ft-lb at 1600 rpm. A map of specific fuel consumption versus bmep is shown in Fig. 24, with volumetric and thermal efficiencies in Fig. 25.

An interesting comparison can be made between the 6-cyl D-282 engine and the 4-cyl 4.125 × 5.250, 281-cu-in. engine which it has replaced in the IH line. Specific fuel consumption versus horsepower at 1800 rpm for these two almost identical displacement engines is shown in Fig. 26. Both engines are the precombustion-chamber type. The improvement in the 6-cyl engine is due completely to lower friction, as shown in Fig. 27, indicating that equivalent combustion efficiencies had been obtained. The lower friction of the D-282 is attributed to shorter, smaller diameter bearings and lower average piston speed. Use of three rings per piston rather than five also contributed to the lower friction. These curves demonstrate that lightweight design can produce benefits beyond lower cost and weight.

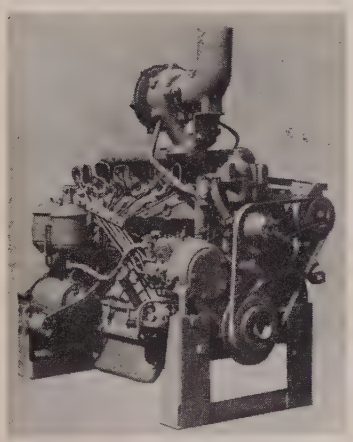


Fig. 28 — DT-282 turbocharged engine

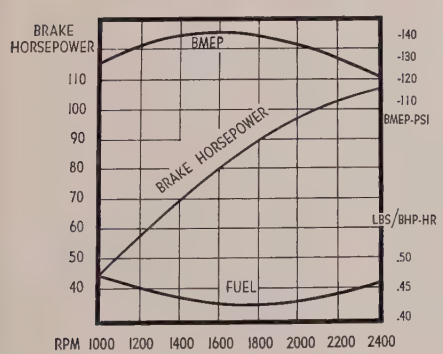


Fig. 29 — UDT-282 turbocharged engine

Turbocharger

It was our natural desire to extend the output range of these low-cost versatile engines still further in order to reach as many applications as possible. The exhaust gas driven turbocharger has been developed for just this purpose, but in 1956 no such units were available for engines of this small displacement. The cost of available units was also out of the question, in view of the cost of the basic engine itself. A brand new turbocharger, designed for low cost and small size, in addition to dependability, was needed. The resulting unit is shown mounted on a D-282 engine in Fig. 28. Its somewhat unusual location at the front of the engine is required to provide clearance in several end-product applications.

The complete turbocharger weighs 24 lb and is capable of continuous operation at 75,000 rpm. Airflows up to 20 lb/min and 2.3 pressure ratio are possible. The alloy iron turbine housing has a single entry with nozzle. The aluminum alloy compressor is of vaneless diffuser design. Elimination of nozzle and vanes provided a broader range of operation, and contributed to the lower cost. The investment cast turbine wheel is brazed to the steel shaft with the unshrouded, uncurved aluminum compressor wheel shrunk onto the same shaft. This rotating

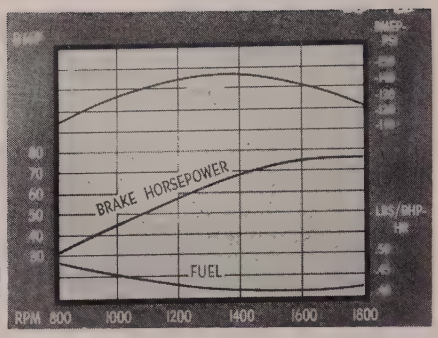


Fig. 30 — Turbocharged DT-282 for TD-9 crawler tractor

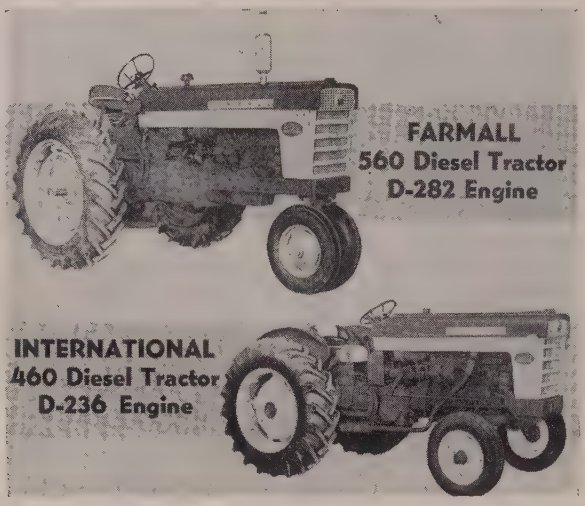


Fig. 31 — Farm tractor applications of IH diesels

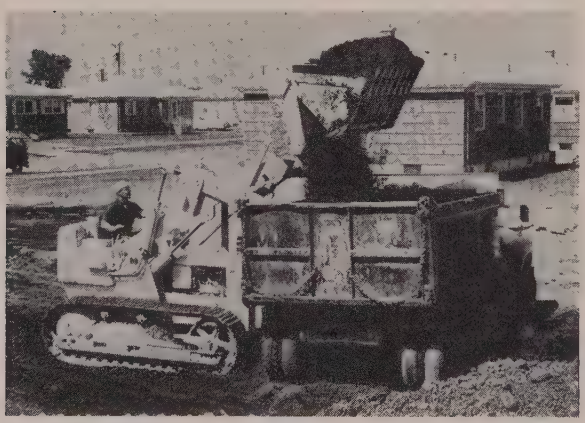


Fig. 32 — TD-9 crawler with DT-282 turbocharged engine

assembly, together with the center bearing housing is designed for field replacement as a package. Cooling and lubrication are provided by oil supplied from the engine lubrication system.

The turbocharged D-282 engine is identical with the naturally aspirated engine except for the special exhaust manifold and turbocharger. Performance characteristics of this engine are shown in Fig. 29. This DT-282 is now used in International's TD-9 crawler tractor, rated 71 net hp at 1800 rpm with a 23% torque rise at 1300 rpm. Horespower, bmep, and fuel economy for this application are shown in Fig. 30.

Fig. 33 — Payloader and stationary applications of IH diesels

All our diesel engines, including this new family, are dynamometer tested prior to shipment to insure the customer of obtaining a well run-in engine which will produce advertised horsepower. Prior to 1954, all production IH diesel engines were tested on block-type electric dynamometers, with power settings made by observing kilowatt output. In 1954, an appropriation of $1,200,000 was used to modernize these production test facilities, located at Melrose Park, Ill., where all IH diesels are built. Sixty-four single-end dynamometers were installed, equipped with dial scales and rpm indicators. In addition, two double-end dynamometers were installed for quality control tests. Fuel rotometers have been recently added.

Engines are run a minimum of 3 hr each. After an initial 2-hr run-in, fuel delivery is set at rated speed and at peak torque speed. Measured torque must be within three pound-feet of specification at both points or pump, nozzle and/or engine are rejected. This procedure has assured consistent engine performance in the field.

Applications

The initial application of these new engines was in IH's farm tractor line announced in July, 1958. Fig. 31 shows the D-282 engine in the Farmall 560, a 5-plow tractor with 69 engine hp. A companion International 560 is available for industrial uses. Fig. 31 also shows the D-236 engine in an International 460 Utility tractor, used for a multitude of farm and industrial uses. The tricycle Farmall 460 with 56 hp also uses the D-236 engine. These applications are all at 1800 engine rpm.

From these initial applications, a wide variety of uses have grown. First was the International 660 tractor, our most powerful wheel tractor using the D-282 engine rated 90 hp at 2400 rpm. Fig. 32 shows the TD-9 crawler tractor using the turbocharged TD-282 engine. Similarly, the TD-6 crawler tractor is now equipped with the naturally aspirated D-282, rated 55 net hp at 1650 rpm. The Model H-50 Hough Payloader, with 5000-lb capacity, uses the D-282 engine, rated 85 hp at 2200 rpm, in conjunction with a torque converter. Many stationary

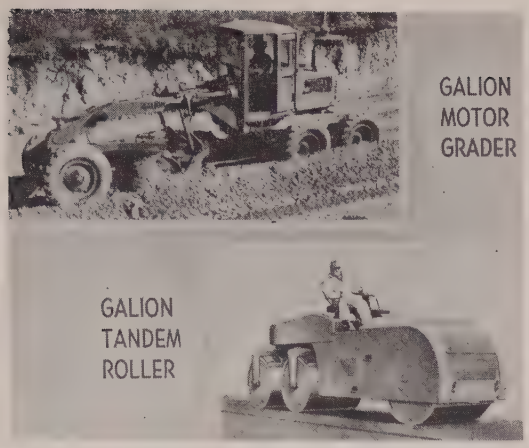


Fig. 34 — Representative OEM applications

applications have been made for irrigation pumps, sawmills, generators and the like (Fig. 33). A multitude of other mobile applications are typified by the Galion motor grader and tandem roller shown in Fig. 34.

D-252 and D-301

The D-252 and D-301 engines were included as members of this new family at the beginning of the paper, but have been specifically excluded from the subsequent detailed discussion. This has been intentional, since these two engines are "late arrivals" and are best compared directly with the D-236 and D-282 from which they grew.

These two engines are intended to extend the versatility of this family in the direction of automotive applications. They result from elimination of the cylinder sleeves, with pistons running directly in the crankcase, producing a 0.125 in. larger bore. Thus, the D-252 is simply a sleeveless D-236, and the D-301 a sleeveless D-282. Only the crankcase bore, pistons, and rings are different. The crankcase castings scheduled for these sleeveless engines are alloyed to improve compatibility between piston and bore.

Most gasoline engines in this type of service have always been sleeveless, and we feel that a sleeveless diesel is completely acceptable. Thousands of hours at 3000 rpm in the laboratory, in proving ground dust at Phoenix, Ariz., and in customer vehicles have supported this viewpoint.

Automotive Applications

The rapidly increasing use of diesel engines in highway vehicles in this country is well-known. Truck registration figures in the "over 26,000 lb" gross vehicle weight class for last year, 1958, show 43% were diesel-powered compared to 29% in 1956. Taxicab installations of diesel engines have become common this year, and the resulting publicity has generated wide public interest. It is clear that the diesel virtues of low fuel consumption and long life between overhauls have not escaped the transportation industry. Only weight and initial cost have prevented wide-spread diesel penetration of this immense market.

The basic design and manufacturing concepts of this IH diesel family have considerably reduced these disadvantages. Not only are these diesels

Fig. 35 — Truck applications: 21,000-lb gvw tractor (left) and ¾-ton pickup (right)

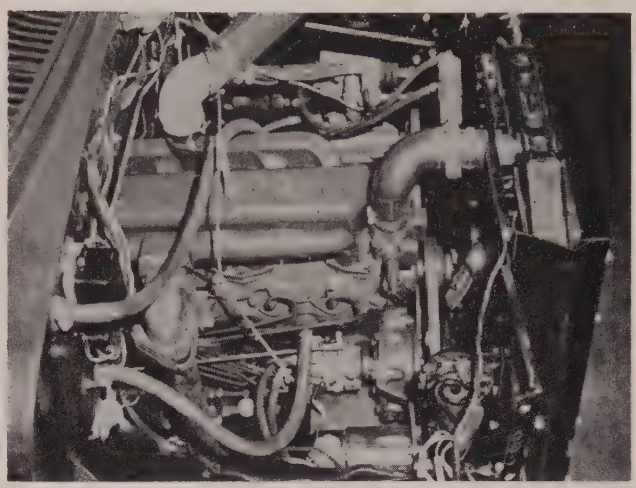


Fig. 36 — D-301 truck installation

lightweight, low cost, and high speed, but they are dimensionally similar to the carbureted engines already used in International trucks (Fig. 35). An evaluation program was begun in early 1958 by installing four modified production engines in test trucks at the IH Motor Truck Engineering Center in Fort Wayne, Ind.

The excellent performance of these engineering vehicles resulted in an extensive field test of trucks equipped with D-301 engines. These engines and trucks were built by our production groups and are now in various customer operations throughout the country. Acceptance of these units has been enthusiastic, and we expect this application to be made generally available soon. Fig. 36 shows the D-301 engine installed in a "cab-forward" truck (upper left in Fig. 35) with its vertical axis inclined 10 deg to the right. This engine develops 110 maximum hp at 3000 rpm and weighs only 860 lb, or 7.8 lb per hp naturally aspirated. This is only 200 lb more than the Black Diamond carbureted engine.

Summary

We feel that our initial goal of lightweight, high-speed diesel engines of lowest possible cost has been more than fulfilled. In fact, applications not originally contemplated for these engines have been successfully made and more are being found. These diesel engines, weighing as little as 7.8 lb. per hp, have been thoroughly proved by over 30,000 hr of laboratory and field tests and by the service of over 24,000 production engines.

Discussion of this paper will be found on p. 512.

DISCUSSION

Bore-Stroke Dimensions Present Interesting Problems

— W. J. Lux
Caterpillar Tractor Co.

THE STORY of the design and development of the four engines just presented is an excellent example of the business-directed type of engineering which is common in our American industry. It might be, and often is, irritating to the engineer to have his management tell him that he must adapt an existing unit, rather than design a new one, or that his design must be machined on existing tooling, or that other apparently arbitrary limits must be observed. Economic considerations are necessary in engineering decisions if the enterprise is to be profitable. The design and development of an engine may be better academically without external restrictions, but the engineer must remember that he is part of a business team.

The choice of engine bore and stroke dimensions presented in this paper is an interesting study. It is not apparent why four engine sizes are needed to cover the horsepower range outlined in Fig. 2. Taking steps proportional to the ratios of the displacements of the four engines, there is a new engine for every 8% interval in rating at both 3000 and 2000 rpm. It would seem that two of these engines, the D-301 and the D-252, for example, could have covered this range quite adequately, having steps of 16% between models. A single model 6-cyl engine and its 4-cyl counterpart would cover this range and an equal lower (or higher) range of rating. In that case, a single piston, valve, rod, bearing, and injection equipment design would be necessary, with the only difference in blocks, heads, crankshafts, and camshafts being the number of cylinders. Balancing in a 4-cyl engine is easily done in a number of ways. The part-load characteristics of the diesel engine certainly allow such a range of ratings, with good performance through that range. Since the external dimensions and major parts of this family of engines are nearly identical, the advantage of the small displacement engine is not apparent in so far as installation space is concerned.

The change in design from the gasoline engine to the dry liner diesel engine appears to be a severe reduction in water passage area between cylinders. Unless the cross-flow of coolant between cylinders is quite low, we would expect some cooling problems in the cylinder area. Doesn't the thin passage also require extra care in the foundry processes at additional expense?

We are interested in this line of engines and the comparative performance of the four models from the standpoint of the different stroke-bore ratios used. Our work in this area, as reported by Kenneth Fleck[a] concluded:

1. Stroke-bore ratios in the 1.25–1.30 range gave higher specific power and lower specific fuel consumption.

2. Torque increase characteristics improved with increased stroke.

The stroke-bore ratios of the four IH Engines are: D-236 — 1.00, D-252 — 0.96, D-282 — 1.19, and D-301 — 1.15. We would expect that the D-282, which was described most completely in the paper, might have the best specific performance of the family.

In an attempt to evaluate these factors, we compared fuel consumption data taken from published specification sheets on the D-282 and the D-236 engines. Fig. A compares the specific fuel consumption at the intermittent ratings for the two engines. The longer stroke D-282 has lower fuel consumption up to the 2400 rpm point, where it appears that the short stroke D-236 is able to take advantage of its ability to run faster. The intermittent ratings for the short stroke engine are also about 5% lower than the displacement difference would indicate they should be. A comparison of minimum specific fuel consumptions (Fig. B) shows the same difference, with the long stroke and lower surface/volume ratio having the apparent advantage. This comparison is of interest to us because it seems to confirm our work on stroke-bore ratios. Are there, perhaps, other factors to explain these performance differences?

The interchangeability of parts between diesel models and between related diesel and gasoline models is clear testimony to thorough planning and good use of existing designs and tooling. Here again, however, we might raise the question of compromising performance (and perhaps space and weight) to obtain such a degree of parts interchangeability. The combustion systems of the four engines, using the same fuel system, camshaft, and other features, but having significantly different main combustion volume shapes, cannot all be optimum. Equally interesting is the use of a common camshaft for both diesel and gasoline models. Here again, it is doubtful that the camshaft timing and profile are optimum for both the diesel and gasoline engines.

The authors have described the details of crankshaft redesign which have resulted in shorter main and connecting-rod bearings. The bearing loads shown in the paper are not out of line: we have good records of long satisfactory service with such loads on steelbacked aluminum bearings. Ratings are shown for these engines up to 3000 rpm. Do the bearing loads shown represent those encountered at 3000 rpm or at the lower tractor rating speeds?

Several references were made to peak pressures of 1800 psi which would be expected in the turbocharged versions of this engine. If ratings are used which result in 1800-psi peak pressures, we wonder if there have been difficulties with cylinder head gaskets on these engines. The increased bores of these engines mean less gasket sealing surface between cylinders (as shown in Fig. 3), this, in turn, places a larger burden upon cylinder head stud spacing, clamping forces between the head and block, and the gasket materials. Do the authors have any comment on this possibility?

We are somewhat uncertain of the meaning of the data on the modified Goodman Diagram in Fig. 13. The diagram is clear if it is assumed that the piston does not rotate in the piston and rod. It is a free floating piston, though, and if it rotates 90 deg, the point which has alternated from zero to the shown maximum tensile stress while it was up at the top, will now alternate between zero and a similar maximum compressive stress while it is at a side position. The range of stress, then, should be approximately doubled for the pin in any of its possible positions. The authors' analysis of the pin stresses by

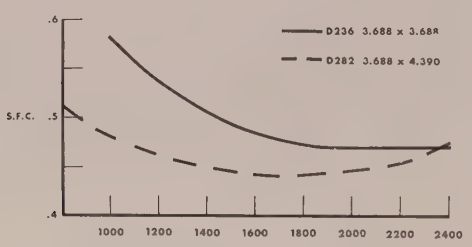


Fig. A — Fuel consumption at intermittent rating

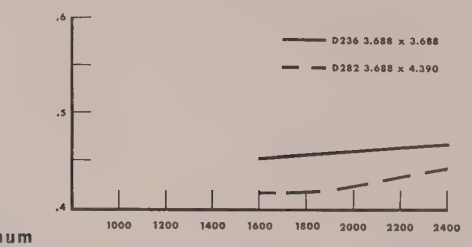


Fig. B — Minimum fuel consumption

[a] "Some Effects of Stroke and Bore on Diesel-Engine Performance," by K. J. Fleck. SAE Transactions, Vol. 65, 1957, pp. 579–594.

way of the thin-walled cylinder is sufficiently accurate for this problem. More complex thick-walled cylinder methods would give a closer answer if a more rigorous analysis were to be made.

The lubricating system arrangement offers some interesting possibilities. The system is actually a parallel combination of a full-flow system and a bypass system; it has perhaps twice as much oil volume passing through it as does a conventional full-flow filter system. Does this affect filter life? Increasing the flow rate through the cooler seems to be a good way to speed oil "warmup" in the engine during cold starts, but it will not result in a lower average oil temperature under full load and speed, even though a larger volume of oil is passing through the cooler.

The speed advance on fuel timing, and the retarded timing for starting are natural developments in high-speed engines. The use of a load advance device is to maintain a nearly constant "start of injection." Was a "constant start of injection pump" considered as a possible way to solve this problem?

The relationship between the D-282 and D-301 and between the D-236 and the D-252 raise another question. If the D-301 and D-252 run on cylinder walls machined into the block casting, would the smaller bore engines do the same thing? The heat-transfer rates through the linear wall should be changed little by the substitution of a thicker cast cylinder wall for the thin-cast wall and a sleeve. The cast cylinder wall could then be bored to either of the two sizes according to automotive practice.

The designs represent another good step forward toward a diesel engine that approaches the physical specifications of a carbureted engine combined with the power and economies of a diesel. The authors point out that the diesel engine is only 200 lb heavier than its gasoline counterpart. This 200 lb, however, is still a 30% increase over the weight of the gasoline engine, and represents the target for future developments in diesel engines.

Authors' Closure To Discussion

MR. LUX raises several questions in his interesting discussion, to which we offer the following answers:

1. The horsepower range is not covered by four engines, but by two engines for industrial use and two (sleeveless) engines for automotive use.

2. No problems have been encountered with coolant flow between the cylinders. It is true that the narrow passages require careful handling at the foundry.

3. Bearing loads listed in Table 1 are for naturally aspirated operation at 1800 rpm.

4. The 1800-psi cylinder pressure was used as a criterion for static testing. Actual turbocharged pressures have not exceeded 1200 psi.

5. Constant start of injection was considered, but available pumps were too large and costly for these engines.

ORAL DISCUSSION

— Reported by Robert J. Dunlap
Murphy Diesel Co.

H. G. Braendel, Wilkening Mfg. Co.: With regard to the piston-ring setup, did the half keystone-type top ring provide better blowby control and/or perform better from a wear standpoint that a full keystone-type ring?

A.: Blowby on the original and conventional ring design was measured at 120 cfm at 1800 rpm full load. Use of the half keystone ring resulted in a reduction in blowby to 60–75 cfm at the same load condition. A full keystone top ring design has never been used and hence, no comparative data are available. Wear with this ring setup has been normal.

Phil Schreiber: Was optimum performance of the diesel members of this family obtained by using valve timing produced by a camshaft which is also used for the carbureted engines in the family? Did you note any detrimental effects on the piston rings when the engine was turbocharged?

A.: Diesel performance was not penalized to the point where we felt justified in designing and producing a camshaft solely for use in the diesels. This was a compromise. Piston-ring blowby is greater on the turbocharged engine. No other problems have arisen. We have not, as yet, tried to compare the naturally aspirated engine with the turbocharged engine in this area.

H. W. Hanners, White Motor Co.: Why is the bmep shown in Fig. 29 higher than that shown in Fig. 30 and how is the torque rise controlled?

A.: Performance curves shown in Fig. 29 are for a stripped engine, while the curves in Fig. 30 are for an engine with accessories. Torque rise is a function of the rate of fuel flow, no turbocharger controls are used.

Burns, Research Center: If the carbureted engines in the family do not require an oil cooler why is one used on the diesel engines?

A.: Laboratory test results for the new diesel engines showed that a lube oil cooler was required.

E. R. Klinge, Ford Motor Co.: Regarding the lube oil cooler requirement, what is the specific heat rejection?

A.: The specific heat rejection is 45 to 50 btu per bhp/min. This includes the heat rejected to the coolant from the oil cooler.

G. R. Beardsley, Ford Motor Co.: Do you have a specific fuel consumption contour map drawn up for the turbocharged engine similar to the one shown in Fig. 24 for the naturally aspirated engine? When will these engines be available for use in truck applications?

A.: We have not drawn up a fuel consumption map for the turbocharged engine. The engines will be available for truck applications in the near future.

R. Lenich, Caterpillar Tractor Co.: Did the original cylinder head design for the diesel engine include the horizontal baffle or was it added later?

A.: The horizontal baffle was a part of the originally designed cylinder head for the diesel members of the family.

G. Wilson, Werner Transportation Co.: Is the oil filter the same one that is used on the Model 264 carbureted engine? What life can be expected from the 301-cu-in. engine in truck applications?

A.: This is not the same filter used on that particular model. Wear life is being evaluated on the basis of several 1000-hr full-load endurance tests in the laboratory. Field tests for wear have not been completed. Initial laboratory tests indicate that wear in the sleeveless engines does not appear to be much different from the sleeved engines.

William Meyer, J. I. Case Co.: What effect does belt tightness have on the damping characteristics of the torsional vibration damper?

A.: No tests have been run to evaluate this. The belt is required under all conditions to drive the water pump.

C. N. Guerasimoff, Allis-Chalmers Mfg. Co.: Why is the oil filter placed behind or downstream from the oil cooler?

A.: Space limitations in the package dimensions for certain applications dictated the location.

A. Larkins: With regard to the crankcase bearing web modifications, what was the reason for placing the diagonal rib in the position as shown in the early experimental engine in Fig. 9?

A.: We don't know the reason for this. The original design was a part of the gasoline engine crankcase design which we inherited.

V. Kuppu Rao, Indian Institute of Science, South India: When the engine is turbocharged, what boost pressure is employed and are injection characteristics, such as rate of injection, changed?

A.: Boost pressure employed in the turbocharged engine is about 20 in. of Hg. Rate of injection for the turbocharged engine was not changed, an increase in the amount of fuel injected is obtained by increasing the length of the injection period.

Reduction of Air Pollution by control of emission from automotive crankcases

P. A. Bennett, C. K. Murphy, M. W. Jackson, and R. A. Randall

Research Laboratories, General Motors Corp.

THE RECENT TECHNICAL literature indicates that air pollution authorities consider the contribution of crankcase gases to the overall hydrocarbon emission from automobiles to be relatively unimportant. Recent reports from the Los Angeles Air Pollution Control District[1] and the Air Pollution Foundation[2] almost completely ignore these gases. Moreover, the proposed vehicle emission standards of the California State Department of Public Health fail to take into account this source. In addition, other workers in the field of automotive air pollution generally have not realized the importance of crankcase emission.[3,4]

One of the few studies that considers this source is a paper by Payne and Sigworth.[5] They published measurements of the concentration of hydrocarbons and oxygenated hydrocarbons, which they defined as "noxious products," in blowby and exhaust gases. They reported that the concentration of "noxious products" in blowby gases was less than three times that in exhaust gases. Considering the relative flow rates of the two gases and these concentration data, it has been generally concluded that crankcase emissions are not a significant source of air pollutants.

The fact that crankcase gases are actually an important source of automobile hydrocarbon emission came to our attention as a result of work in another field. During the development of an engine performance test for evaluating automotive crankcase lubricants, a definite correlation between intake air humidity and engine rusting tendency was observed.[6] Consideration of these observations led to the conclusion that blowby gases are predominately carbureted fuel-air mixture. This work suggested that blowby and crankcase vent gas hydrocarbon concentrations were actually some 30 times (rather than three times) that of exhaust gas.

This paper: (1) reports recent data on the composition of blowby gases, (2) evaluates the relative importance of crankcase and exhaust gas hydrocarbon emissions using the best available analytical techniques, and (3) presents and evaluates a method of eliminating the contribution of automotive crankcase gases to air pollution.

ANALYSES of crankcase vent gases by several methods have shown that crankcase and exhaust hydrocarbon emissions from automobiles are of the same order of magnitude.

Internal ventilation of the crankcase to the engine intake system eliminates crankcase emission, thus providing a practical control of this important source of air pollution.*

Equipment and Procedures

Several of the terms used in this paper are defined in the schematic diagram shown in Fig. 1. Blowby gases consist of a mixture of unburned fuel-air charge and exhaust products blown by the piston rings. Crankcase vent gases are composed of these blowby gases plus ventilation air. During normal car operation, ventilation air drawn in through the

* Paper presented at SAE Annual Meeting, Detroit, Jan. 14, 1960.

[1] "Total Air Pollution Emissions in Los Angeles County," by R. L. Chass, R. C. Lunche, N. R. Schaffer, and P. S. Tow. Paper presented at Air Pollution Control Assoc. 52nd Annual Meeting, Los Angeles, June, 1959.

[2] "Auto Exhaust in Air Pollution," by W. L. Faith. Paper presented at American Chemical Society meeting, Atlantic City, September, 1959.

[3] "Smogless Automobile," by F. L. Hartley, C. C. Moore, and J. B. Gregory. Paper presented at Air Pollution Control Assoc. 52nd Annual Meeting, Los Angeles, June, 1959.

[4] "Carburetor Evaporation Losses," by J. T. Wentworth. Paper presented at SAE Annual Meeting, Detroit, January, 1958.

[5] "Composition and Nature of Blowby and Exhaust Gases from Passenger-Car Engines," by J. Q. Payne and H. W. Sigworth. Proceedings of Second National Air Pollution Symposium, Vol. 2, 1952, p. 62.

[6] "Engine Rusting and General Motors MS Test," by P. A. Bennett. Paper presented at National ASLE Meeting, Cincinnati, April, 1960.

breather is mixed with the blowby gases, and the resulting mixture flows out through the road draft tube. It should be noted that when the breather and dip stick tube are plugged, the gases coming out of the draft tube are blowby gases. This definition ignores the absorption of blowby constituents by the lubricants or any inclusion of entrained or vaporized lubricants. A detailed description of the equipment and procedures used is presented in Appendix I.

The composition of exhaust and crankcase vent or blowby gases was determined by: (1) Orsat analyses of grab samples obtained before and after a combustion train, (2) gas chromatographic analyses of grab samples, and (3) continuous infrared analyses.

When using the combustion train-Orsat technique, the breather and dipstick tube were plugged. The draft tube was restricted to pressurize the crankcase and prevent air dilution of these gases. The pressurized blowby gases were forced through a heated catalyst tube containing Hopcalite to oxidize the combustibles. Grab samples were collected before and after the catalyst tube and subsequently analyzed by Orsat techniques for carbon dioxide, oxygen, and carbon monoxide. Exhaust gas samples were analyzed by a similar technique.

In order to determine the individual hydrocarbon concentrations and the absolute total hydrocarbon concentrations, samples of the crankcase vent or blowby gases were collected and subsequently analyzed by gas chromatography. Samples of the fuels used in these tests were also analyzed by gas chromatography.

Exhaust and crankcase vent or blowby hydrocarbon concentrations were measured continuously with a Beckman L/B Model 15A infrared analyzer. The analyzer was equipped with a n-hexane detector and quartz optics. Due to the inability of the analyzer to respond equally to all hydrocarbon types, the values determined with the analyzer are not absolute concentrations. However, through the use of gas chromatography, the measured Beckman concentrations were converted to absolute hydrocarbon concentrations.

Exhaust gas flow rates were calculated from measured fuel flow rates and Orsat-determined air/fuel ratios. Crankcase vent or blowby flow rates were determined by a tracer gas technique based on the following observation; chromatographic analyses of numerous samples of exhaust and crankcase gases have shown that these gases contain less than 1 ppm propane. Consequently, propane as the tracer was metered into the crankcase. Grab samples of the crankcase vent or blowby gases were analyzed for propane by gas chromatography. From the propane concentrations and metered flow rates, the crankcase vent or blowby flow rates were calculated.

Most tests were conducted on the road at selected car speeds from 0 to 60 mph and various throttle settings. A few tests were conducted on an engine run on a dynamometer.

Blowby Composition

The composition of blowby gases is an important factor in determining the magnitude of automotive crankcase emission. Typical Orsat analyses of exhaust and blowby gases are shown in Table 1. For both the rich and lean air/fuel ratios, the blowby

Table 1 — Orsat Analyses of Exhaust and Blowby Gases
(Engine of Make Z — Dynamometer Tests)

	Volume Per Cent		
	CO_2	O_2	CO
Rich Air/Fuel (12.6)			
Exhaust Gas	11.0	0.7	4.6
Blowby Gas	1.5	17.9	0.4
Lean Air/Fuel (16.4)			
Exhaust Gas	13.0	2.6	0.5
Blowby Gas	1.8	18.4	0.3

Table 2 — Effect of Oxidation on Orsat Analyses of Exhaust and Blowby Gases
(Engine of Make Z — Dynamometer Tests)

	Volume Per Cent					
	Normal			Oxidized		
	CO_2	O_2	CO	CO_2	O_2	CO
Lean Air/Fuel (16.4)						
Exhaust Gas	13.0	2.6	0.5	13.5	2.1	0.0
Blowby Gas	1.8	18.4	0.3	11.9	4.0	0.0

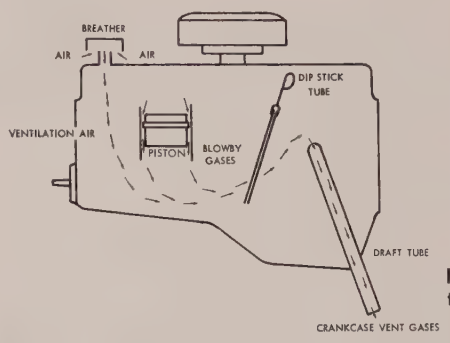


Fig. 1 — Conventional crankcase ventilation system

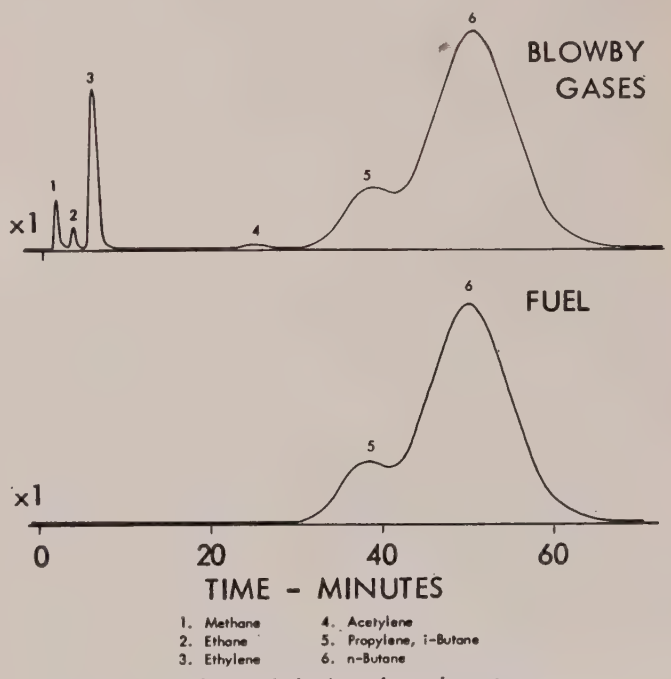
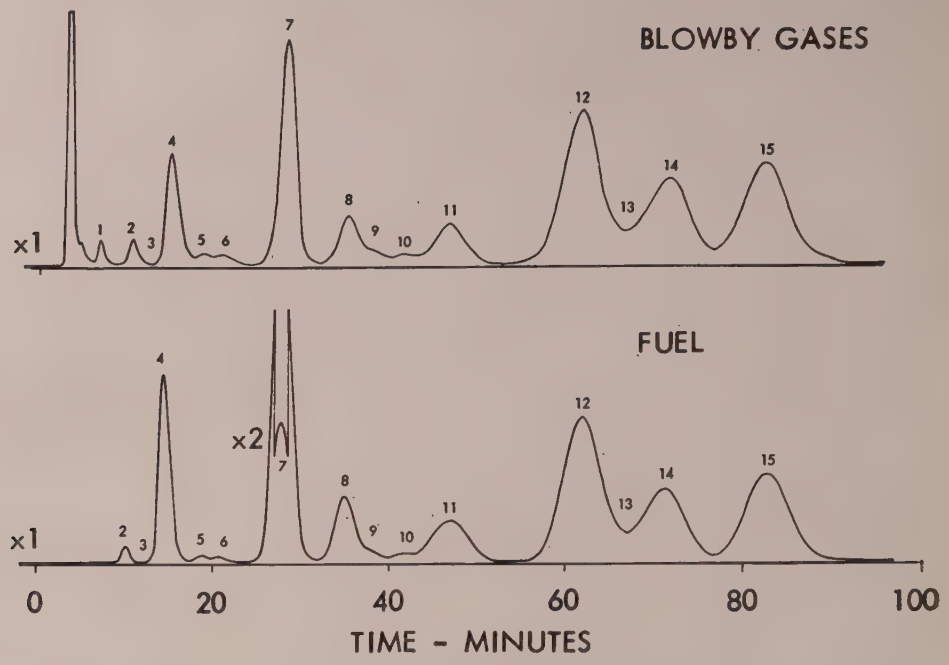
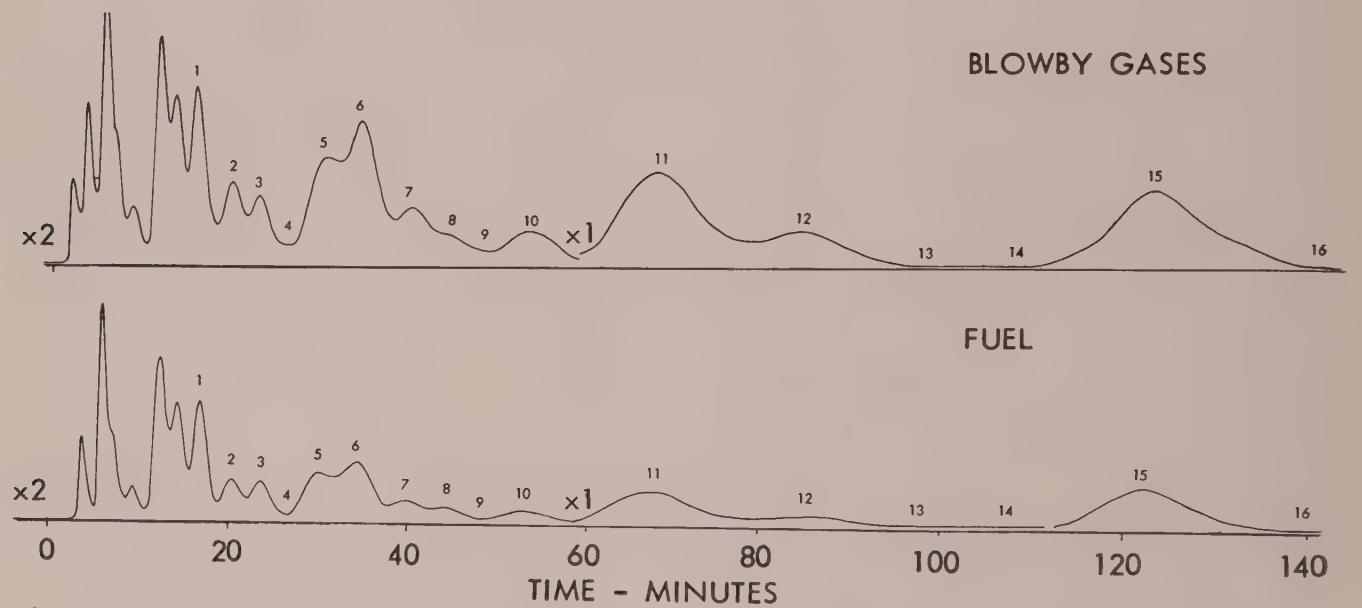


Fig. 2 — Silica gel-alumina column chromatograms



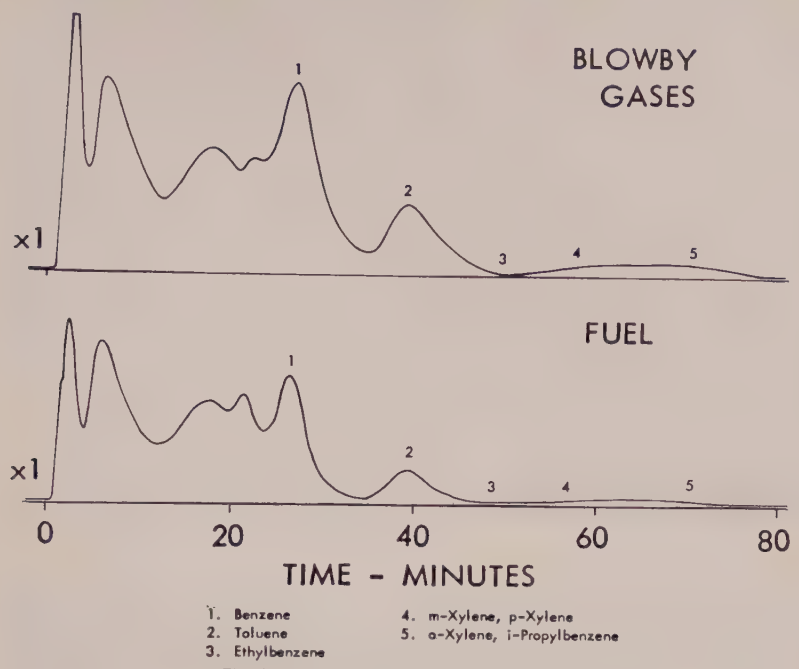
1. Propylene
2. i-Butane
3. Methylacetylene, Propadiene
4. n-Butane, 1-Butene, i-Butene
5. 2-Butene (Trans), 1,3-Butadiene
6. 2-Butene (Cis)
7. i-Pentane, 3-Methyl-1-butene
8. n-Pentane, 1-Pentene
9. 2-Methyl-1-butene
10. 2-Pentene, 2,2-Dimethyl-3-butene
11. 2-Methyl-2-butene, 2,2-Dimethylbutane, 2-Methyl-1,3-butadiene
12. 2-Methylpentane, 2,2-Dimethylbutane, 2-Methyl-4-pentene
13. Cyclopentane, 2,3-Dimethyl-1-butene, 2-Methyl-3-pentene
14. 3-Methylpentane
15. n-Hexane, 1-Hexene, 2-Methyl-1-pentene

Fig. 3 — Phthalate column chromatograms



1. n-Hexane, 1-Hexene, 2-Methyl-1-pentene
2. 3-Ethyl-1-butene, 2-Hexene, 3-Hexene, 2,2-Dimethylpentane, 2-Methyl-2-pentene 2,4-Dimethylpentane
3. Methylcyclopentane, 3-Methyl-2-pentene, 2,3,3-Trimethylbutane, 2,3,-Trimethyl-1-butene 2,2-Dimethyl-2-butene, 3,4-Dimethyl-1-pentene
4. 3,3-Dimethylpentane
5. 2,3-Dimethylpentane, 5-Methyl-1-hexene, 3,4-Dimethyl-2-pentene, Cyclohexane, 3-Methylhexane
6. 2,4,4-Trimethylpentane, 3-Ethylpentane, 5-Methyl-2-hexene
7. n-Heptane, 1-Heptene, Cyclohexene
8. Benzene
9. 2,4,4-Trimethyl-1-pentene, 3-Heptene
10. Methylcyclohexane, 2,4,4-Trimethyl-2-pentene, 2,2,3-Trimethylpentane
11. 2,3,4-Trimethylpentane
12. 3-Methylheptane, 2,2,5-Trimethylhexane
13. n-Octane
14. 1-Methylcyclohexene, 2-Ethyl-1-hexene, 1-Octene
15. Toluene, 2,3-Dimethyl-2-hexene, 2,6-Dimethyl-3-heptene, 2-Octene (Cis)
16. 2-Octene (Trans)

Fig. 4 — Silicone column chromatograms



1. Benzene
2. Toluene
3. Ethylbenzene
4. m-Xylene, p-Xylene
5. o-Xylene, i-Propylbenzene

Fig. 5 — Aroclor column chromatograms

Table 3 — Sample Data and Calculations
(Car A — 30 mph — Road Load)

Crankcase Vent Gases	
(a) Beckman Hydrocarbon Concentration	6700 ppm C_6
(b) Absolute Hydrocarbon Concentration, (1.25) (a)	8250 ppm C_6
(c) Flow Rate	1.4 cfm
(d) Hydrocarbon Flow Rate, (b) (c)	0.0115 cfm C_6
(e) Hydrocarbon Emission Rate, (d) (0.206 lb C_6H_{14}/cu ft) (60 min/hr)	0.14 lb C_6/hr
Exhaust Gases	
(a) Beckman Hydrocarbon Concentration	250 ppm C_6
(b) Absolute Hydrocarbon Concentration, (1.8) (a)	450 ppm C_6
(c) Flow Rate	26 cfm
(d) Hydrocarbon Flow Rate, (b) (c)	0.0117 cfm C_6
(e) Hydrocarbon Emission Rate, (d) (0.206 lb C_6H_{14}/cu ft) (60 min/hr)	0.14 lb C_6/hr

gases had relatively low concentrations of carbon dioxide and carbon monoxide and high oxygen concentrations. Because the crankcase was pressurized to avoid air dilution of the sample, the high oxygen concentration could not have resulted from air leakage. Therefore, the oxygen must have entered the crankcase in the gases blown by the piston rings. Because exhaust gases contain low concentrations of oxygen, the oxygen present in the blowby gases must come from the carbureted mixture. Furthermore, since the high oxygen concentration indicates that blowby is predominately carbureted mixture, the gases should contain large amounts of hydrocarbon.

Similar findings have been reported by Cutter[7] who stated in 1933, "It can be unquestionably concluded . . . that a large part of blowby is unburned charge." This conclusion has been inferred more recently by Spindt, Wolfe, and Stevens.[8]

In order to determine the amount of hydrocarbons present in the blowby gases, these gases were burned in a combustion train using a Hopcalite catalyst. Samples collected before and after the combustion train were analyzed by Orsat techniques. The data in Table 2 show that oxidation of blowby gases resulted in a large increase in the carbon dioxide concentration. Oxygen concentration decreased markedly. Consequently, the combustion train-Orsat analyses established that the blowby gases contain large amounts of combustible material. Exhaust gases were oxidized by a similar technique and the results indicated that the combustible concentration was small.

In order to identify the combustibles in blowby gases, grab samples were collected and analyzed by gas chromatography. The results of a typical blowby gas analysis are given in Figs. 2–5. The top chromatogram in each figure is that of the blowby gases; the lower chromatogram is that of the fuel used in this test. In Figs. 3–5 the components in the first few unidentified peaks result from unseparated carbon monoxide and hydrocarbons. The hydrocarbons are separated in one of the other columns and are designated on the respective figure. It should be noted that equal area peaks from different columns do not represent equal hydrocarbon concentrations because of different column flow rates and sample sizes. Note the similarity between the fuel and blowby chromatograms from the phthalate, silicone, and aroclor columns (Figs. 3–5). Only the silica gel-alumina column shows pronounced differences between the fuel and blowby chromatograms. The small methane, ethane, and ethylene peaks of the blowby chromatogram are completely absent in the fuel analysis. Alternately, exhaust products contain large amounts of these materials.[5,9] The presence of cracked products in the blowby is due to the exhaust gases present. These chromatograms indicate that blowby hydrocarbons are predominately fuel, a conclusion which is supported by the data of Payne and Sigworth.[5]

It can be estimated from the Orsat analyses of blowby and exhaust gases (Table 1) that blowby is approximately 85% carbureted mixture. The remaining 15% is exhaust products. By assuming that carbureted mixture has a hydrocarbon concentration of 20,000 ppm C_6 and that exhaust has a concentration of 400 ppm C_6, it can be further estimated that less than 1% of the hydrocarbons present in blowby gases would be from the exhaust fraction. Over 99% would be directly from the carbureted fuel. Accordingly, specific hydrocarbons in the fuel will determine the specific hydrocarbons in the

[7] "Study of Blowby in Internal Combustion Engine," by D. C. Cutter. Unpublished thesis for Degree of Mechanical Engineer, Yale University, 1933.
[8] "Nitrogen Oxides, Combustion, and Engine Deposits," by R. S. Spindt, C. L. Wolfe, and D. R. Stevens. SAE Transactions, Vol. 64, 1956, pp. 797–811.
[9] "Exhaust Gas Analysis by Gas Chromatography Combined with Infrared Detection," by W. B. Heaton and J. T. Wentworth. *Analytical Chemistry*, Vol. 31, 1959, p. 349.

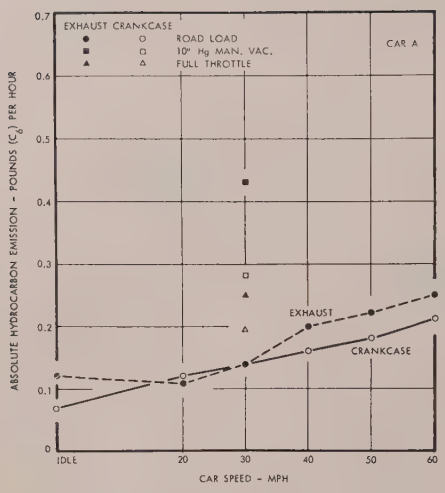


Fig. 6 — Crankcase and exhaust hydrocarbon emission, Car A

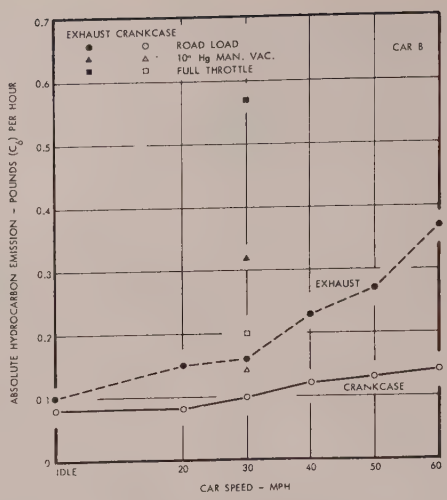


Fig. 7 — Crankcase and exhaust hydrocarbon emission, Car B

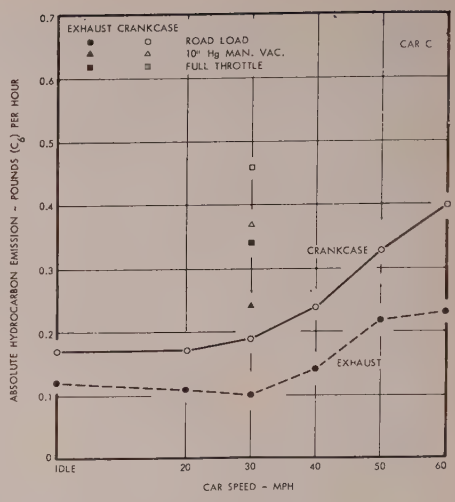


Fig. 8 — Crankcase and exhaust hydrocarbon emission, Car C

crankcase vent gases. This conclusion is supported by the chromatographic analyses of fuel and blowby gases (Figs. 2–5). The variation in blowby gas composition with fuel type determined with a mass spectrometer[5] further substantiates this conclusion.

A number of factors which help to explain the high hydrocarbon content of engine blowby are discussed in Appendix II. A consideration of the effects of flame travel and flame quenching offer a plausible explanation for the composition of these gases.

Crankcase Emissions

To cover the range of results observed in a number of studies, the levels of crankcase hydrocarbon emission from five cars on the road will be presented and discussed. Results from three 1959 cars having widely different crankcase hydrocarbon emission rates are presented. Also included are the results from one 1950 car with high oil consumption and one 1955 car with high blowby rates as well as high oil consumption. Complete data for the five cars are listed in tabular form in Appendix III. The results for each car will be discussed individually.

Car A — To illustrate the calculations and assumptions needed to arrive at emission rates, a portion of the data for car A at 30 mph-road load is presented in Table 3. For the crankcase gases, the Beckman hydrocarbon concentration of 6700 ppm is multiplied by 1.25 (Appendix I) to give an absolute hydrocarbon concentration of 8250 ppm. Multiplication by the crankcase vent gas flow rate of 1.4 cfm gives a hydrocarbon flow rate of 0.0115 cfm. Assuming the hydrocarbon is C_6H_{14} and that the perfect gas law applies, results in a hydrocarbon gas density of 0.206 lb/cu ft at 90 F and 1 atm. This density, along with a conversion to an hourly basis, gives a crankcase hydrocarbon emission rate of 0.14 lb/hr.

The exhaust gas numbers are treated similarly; however, the Beckman concentration of 250 ppm is multiplied by a factor of 1.8 (Appendix I) to obtain an absolute hydrocarbon concentration of 450 ppm.

Table 3 also compares hydrocarbon concentrations and flow rates of the exhaust and crankcase vent gases. The hydrocarbon concentration of crankcase vent gas is about 20 times that of exhaust gas. On the other hand, the exhaust flow rate is about 20 times the crankcase ventilation rate. The net effect is that the crankcase and exhaust have equal hydrocarbon emission rates.

The hydrocarbon emission rate of Car A in pounds per hour is plotted as a function of car speed in Fig. 6. The road load points are connected by lines. Two heavier load conditions, 10 in. of Hg manifold vacuum and full throttle are denoted by other symbols.

The data show that the crankcase and exhaust hydrocarbon emission rates were about the same for each speed and load condition studied. This general relationship of crankcase to exhaust hydrocarbon emission is about the average for all cars which have been studied. The crankcase hydrocarbon emission rate for Car A ranged from a low of 60% of the exhaust hydrocarbon rate at idle to a high of 110% at 20 mph-road load.

Car B — In a similar manner, Fig. 7 summarizes the hydrocarbon emission rates for Car B. Not only were the crankcase emission rates lower than the rates for Car A, but the exhaust hydrocarbon emission rates were higher than for Car A. As a result the crankcase hydrocarbon emission rates are low relative to exhaust emission rates. The crankcase emissions ranged from a low of 35% of exhaust emission at 30 mph-full throttle to a high of 80% at idle.

Car C — The results for Car C (Fig. 8) show the opposite trend from Car B. The exhaust hydrocarbon emission rate was nearly the same as for Car A. However, the crankcase emission rate was nearly double that of Car A at all conditions. This resulted in the crankcase hydrocarbon emission rates of Car C always being much greater than the exhaust emission rates. Crankcase emissions were as low as 140% of exhaust emission rates at idle and as high as 190% at 30 mph-road load.

No explanation for the high crankcase hydrocarbon emission rates of Car C is possible at this time. This 1959 car had no obvious malfunctions and had normal oil consumption (2000 mpq).

Car D — Car D was included to study the relationship between oil consumption (about 150 mpq) and crankcase hydrocarbon emission rates. The

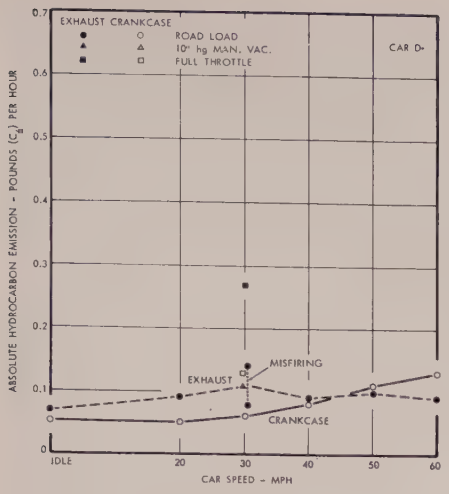


Fig. 9 — Crankcase and exhaust hydrocarbon emission, Car D

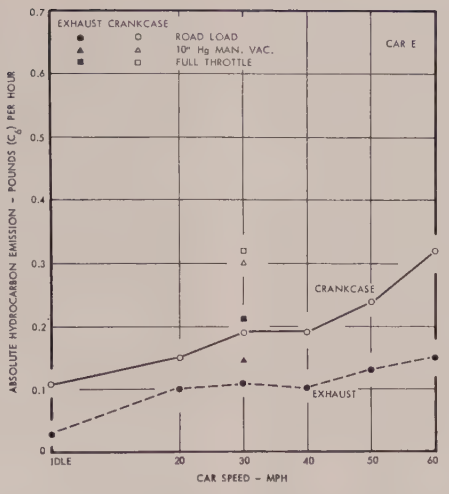


Fig. 10 — Crankcase and exhaust hydrocarbon emission, Car E

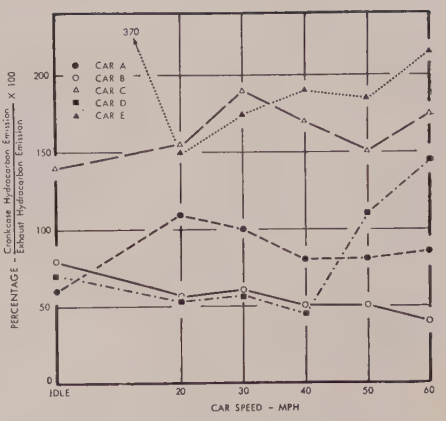


Fig. 11 — Crankcase hydrocarbon emission as percentage of exhaust hydrocarbon emission

data for this older car are summarized in Fig. 9. Because the breather was plugged to ensure that no backflow existed, the crankcase gases are considered to be blowby.

Both the exhaust and blowby hydrocarbon emission rates were low for this car. The low crankcase emission rates result from low blowby flow rates. Therefore, the high oil consumption of this car was not an indication of high crankcase hydrocarbon emission rates. The crankcase hydrocarbon emission rate ranged from a low of 45% of the exhaust hydrocarbon emission rate to a high of 145%.

Car E — Although Car D had high oil consumption, it did not prove to have the expected high blowby rates. Car E, however, had high blowby flow rates as well as high oil consumption (about 50 mpq). As with Car D, the breather was plugged for the tests. Results for Car E are presented in Fig. 10.

The crankcase hydrocarbon emission rate was as high as 370% of the exhaust hydrocarbon emission rate at idle and as low as 150% at 20 mph-road load. These high crankcase emission rates result from the high blowby flow rates (Table 11, Appendix III). The results for Cars D and E show that crankcase hydrocarbon emission rates cannot be predicted from age or oil consumption. Although the blowby hydrocarbon concentration of Car D was different from that of Car E, the blowby flow rate appears to be a good indication of the magnitude of crankcase hydrocarbon emission (Tables 10 and 11). Despite the poor mechanical condition of the engines in Cars D and E, the exhaust hydrocarbon emission rates were low.

Comparison of Crankcase and Exhaust Hydrocarbon Emissions — A summary of the crankcase and exhaust hydrocarbon emission results for the five cars at road load is presented in Fig. 11. The crankcase emission rates are expressed as a percentage of the exhaust emission rates. A value of 100% denotes equal exhaust and crankcase emission rates. The lowest value is 40%; the highest is 370%. Although there is a wide range in the percentages, it can be concluded for these cars under these road load conditions that crankcase and exhaust hydrocarbon emission rates are of the same order of magnitude.

To compare the relative contribution to air pollution by crankcase and exhaust gases, it is necessary to consider emissions during all operating conditions. For idle, road load, and accelerating conditions, emission rates are similar from both crankcase and exhaust. On the other hand, although exhaust hydrocarbon emissions are appreciable during deceleration, crankcase emissions are negligible because deceleration blowby rates are essentially zero.

If it is assumed that 30% of the total exhaust hydrocarbon emission occurs during deceleration, the following estimate can be made. Deceleration emissions account for 30% of the exhaust emissions; idle, road load, and acceleration emissions account for the other 70%. Crankcase emissions — equal to idle, road load, and acceleration exhaust emissions — are equivalent to 70% of the exhaust emission. Converting to a total emission basis, crankcase hydrocarbon emissions amount to 70/170 or 40% of the entire engine hydrocarbon emission (exclusive of carburetor vent losses). Therefore, it must be concluded that crankcase hydrocarbon emission is a major portion of the automobile's contribution to air pollution.

Control of Crankcase Emissions

One way to eliminate hydrocarbon emission from crankcases is to feed the crankcase gases back into the induction system of the engine. The introduction of crankcase gases into the intake system in truck engines has been used to improve crankcase cleanliness.[10] It is clear that this system would eliminate the direct emission of crankcase gases, but testing was required to ensure that this system did not increase the hydrocarbon emission of exhaust gases. Therefore, a testing program was conducted.

Equipment — A sketch of an internal crankcase ventilating system is shown in Fig. 12. The road draft tube is replaced by an adaptor, a vent valve,

[10] "Motor Oils and Engine Lubrication," by C. W. Georgi. Pub. by Reinhold Publishing Corp., New York, 1950.

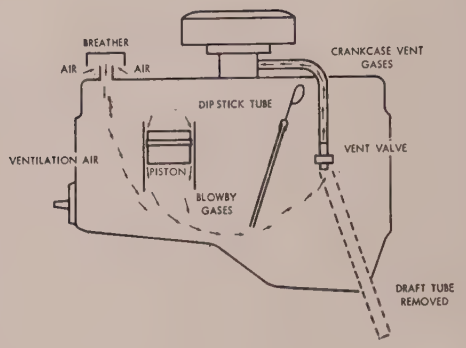


Fig. 12 — Internal crankcase ventilation system

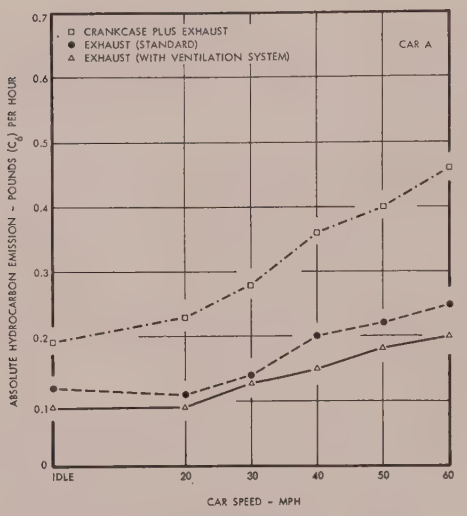


Fig. 13 — Hydrocarbon emission with and without internal crankcase ventilation, Car A

and a tube connected to the intake system below the carburetor.

One of the most important parts of the internal crankcase ventilation system is the vent valve. This valve must have sufficient capacity to handle the blowby gases plus a required amount of crankcase ventilating air. However, it must control the flow so that carburetion is not upset, causing misfiring which increases exhaust hydrocarbon emission. In addition, the valve must not permit backfiring into the crankcase.

The vent valves used in this work controlled the flow through the ventilation system in the following way. At idle and road load manifold vacuums the valve maintained the flow between 1 and 2 cfm. At intake manifold vacuums between 14 and about 2 in. of Hg, the valve permitted flow rates up to 5 or 6 cfm. Flow was reduced at full throttle by movement of the valve toward a seating position which was incorporated to prevent backfiring into the crankcase.

The internal crankcase ventilation system is capable of handling blowby gases under most operating conditions without upsetting carburetion. With an engine in good mechanical condition, momentary backflow of crankcase gases out the breather may occur at full-throttle conditions which are infrequently encountered in urban areas. With engines having excessive blowby, the internal crankcase ventilating system may not be entirely effective.

With the assistance of representatives of the car manufacturing companies, the particular internal crankcase ventilating system for each car was selected. The systems used are not necessarily the exact systems which would now be chosen for commercial application, because of additional development and testing since these results were obtained. However, the trends should be indicative of results obtainable by ventilation of the crankcase into the engine intake.

After installation of the crankcase ventilating system, idle adjustments were made and measurements of exhaust hydrocarbon concentrations and flow rates were made. Using the tracer gas technique, checks were made at several critical conditions (high speed-road load and full throttle) for

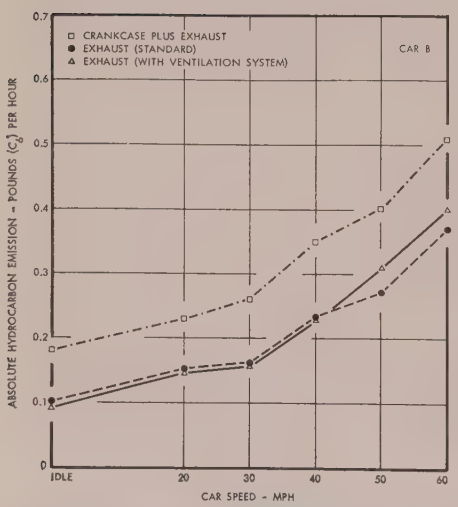


Fig. 14 — Hydrocarbon emission with and without internal crankcase ventilation, Car B

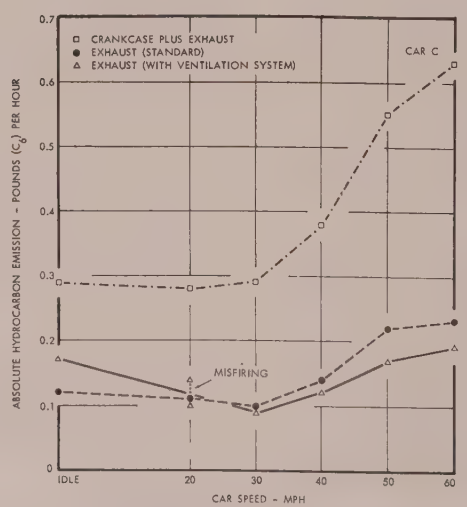


Fig. 15 — Hydrocarbon emission with and without internal crankcase ventilation, Car C

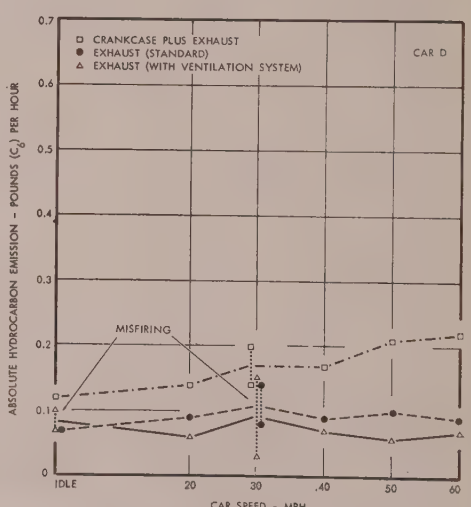


Fig. 16 — Hydrocarbon emission with and without internal crankcase ventilation, Car D

backflow out of the breather.

The results obtained on the five cars with internal crankcase ventilation are presented along with the standard car emission results in Appendix III.

Car A — The results obtained after the installation of an internal crankcase ventilation system on Car A are shown in Fig. 13. The method of presentation is similar to that used in the preceding figures. For clarity, the high load data are omitted, but are listed in Table 7. The high load data follow the general pattern shown for the road load conditions.

The upper curve results from a summation of the crankcase plus exhaust hydrocarbon emissions of the standard engine. This is the total emission of the engine (exclusive of carburetor losses) replotted from the curves in Fig. 6. The center curve is the exhaust hydrocarbon emissions from the standard car also replotted from Fig. 6. The lower curve is the exhaust emission with the internal crankcase ventilation system installed. No backflow was detected at these conditions; thus, the lower curve represents the total hydrocarbon emission from the engine (exclusive of carburetor losses) when the crankcase gases are fed to the intake system. With this car, the resulting emission appears to be lower than the standard engine exhaust emission. However, at most speeds the difference may not be significant.

Car B — Similarly, the results obtained with Car B are presented in Fig. 14. Installation of the internal crankcase ventilation system eliminated the crankcase hydrocarbon emission without appreciably affecting exhaust hydrocarbon emissions. As with Car A, measurements showed no backflow out of the breather at any condition.

Car C — The hydrocarbon emission results for Car C with and without the crankcase vented to the intake are shown in Fig. 15. In this V-8 engine with 4-barrel carburetion, the crankcase gases were introduced in a different position in the intake system from that used in the other cars. In this car the gases were introduced into one-half of the intake manifold; in the other cars introduction was immediately below the carburetor.

As with Cars A and B, the exhaust emission was little affected by internal crankcase ventilation. Backflow did not occur at any of the conditions shown on the figure. However, backflow was detected at 30 mph-full throttle. At this condition, which is infrequently encountered in urban driving, crankcase gases in excess of the flow capacity of the ventilating system flowed out of the breather.

In the figure the variation in exhaust hydrocarbon emissions at 20 mph with internal crankcase ventilation was attributed to misfiring. This misfiring illustrates one of the problems which accompany the installation of such a system. Carburetor idle adjustment was very critical, and it was impossible to adjust the idle to eliminate all misfiring. The ventilation system was modified to introduce crankcase gases into both sides of the manifold; no evidence of misfiring was noted.

A car of the same make with a 2-barrel carburetor was run with and without crankcase gases vented into one-half of the intake manifold. No evidence of misfiring was found at any condition. Therefore, this engine-carburetor combination was less sensitive than the combination in Car C to the introduction of crankcase gases into one-half of the intake manifold.

Car D — Fig. 16 presents hydrocarbon emission results for Car D with and without internal crankcase ventilation. Venting the crankcase gases back to the engine intake had no appreciable effect on exhaust hydrocarbon emission rates. The many points at 30 mph result from misfiring which was present in the car with or without the ventilating system. Installation of the ventilating system also resulted in misfiring at idle. No attempt was made to eliminate the misfiring by special adjustment or repairs.

Car E — The results obtained with Car E are presented in Fig. 17. As discussed earlier, this particular car had excessively high blowby flow rates. Therefore, the results presented for Car E differ from the results presented on the other cars, in that the internal crankcase ventilating system used could not handle the blowby flow rate even at road load conditions. From the blowby flow rates on this car (Table 11) and the flow capacity of the vent valve (1.5 cfm at road load manifold vacuums), it was possible to calculate the crankcase hydrocarbons which backflow out of the breather. The middle curve on Fig. 17 results from a summation of the "breather" hydrocarbons and exhaust hydrocarbons obtained with the internal crankcase ventilation system.

The excessive blowby flow rates resulted from the poor mechanical condition of this engine. Subsequent inspection showed that the piston rings were very badly worn. Because the blowby flow rate exceeded the capacity of the ventilation system, only about 50% of the hydrocarbons emitted from the crankcase were eliminated. With the ventilation system used, any increase in the flow capacity to handle these excessive blowby rates would result in carburetion difficulties. The mechanical condition of this engine dictates a major overhaul to reduce the excessive blowby flow rates.

Effectiveness of Internal Crankcase Ventilation — The effectiveness of internal crankcase ventilation systems on the five cars in reducing engine hydrocarbon emissions is summarized in Fig. 18. The reduction of hydrocarbon emission is expressed as a

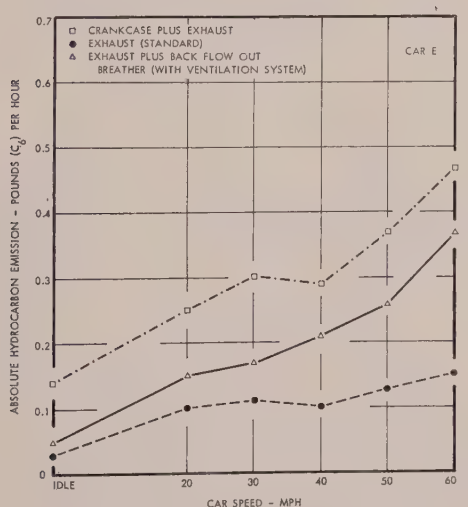


Fig. 17 — Hydrocarbon emission with and without internal crankcase ventilation, Car E

percentage of the exhaust plus crankcase hydrocarbons eliminated by the ventilation system. The greater the percentage reduction, the greater the effectiveness. The values for Cars A, C, and D were between 50 and 70%. Car B, which had low crankcase hydrocarbon emission relative to the exhaust, had reductions between 25 and 50%. The reductions for Car E were also low, especially at the higher speeds, because the blowby flow rates exceeded the flow capacity of the internal ventilation system.

Summary and Conclusions

From the measurements of hydrocarbon concentrations and flow rates, it has been established that crankcase and exhaust hydrocarbon emissions are of the same order of magnitude. Therefore, it must be concluded that crankcase hydrocarbon emission is a major portion of the automobile's contribution to air pollution.

Ventilation of the engine crankcase to the intake system eliminated crankcase hydrocarbon emissions without appreciably affecting exhaust emissions. However, a successful internal crankcase ventilation system must be carefully engineered for each engine design with particular regard to carburetion and manifolding.

Considering the crankcase and exhaust hydrocarbon emissions under all operating conditions — idle, road load, acceleration, and deceleration — it was estimated that the internal crankcase ventilating system eliminates 40% of the engine hydrocarbon emissions exclusive of carburetor vent losses.

Various analyses have led to the conclusion that blowby gases are predominately carbureted mixture. This indicates that the specific hydrocarbons in the fuel determine the specific hydrocarbons in the crankcase gases. Therefore, in cars in which the crankcase emissions are not controlled, the fuel composition determines the crankcase hydrocarbons emitted.

Despite the fact that the importance of crankcase emission was only recently recognized, the crankcase has been, and continues to be, a major source of hydrocarbons relative to exhaust. Any control of air pollution from the automobile must be formulated with this fact in mind.

Acknowledgments

The authors would like to thank members of the Fuels and Lubricants Department for their valuable advice and assistance, especially that of Thomas Prater in obtaining the chromatographic data, and to express their appreciation to the General Motors employees whose personal cars were used in this investigation.

APPENDIX I

Equipment and Procedures

Sampling System — The sampling system consisted of either a tail pipe probe or a draft tube probe, water trap, filter, rotameter, analyzer, and a pump. A schematic diagram of the sampling system is shown in Fig. 19. The probe was inserted at least 2 ft into the tail pipe or the road draft tube.

The flow rate through the sampling system was controlled by Valve A which bleeds air to the pump. Flow rates through the analyzer of 0.3 and 0.03 cfm were used for the exhaust and crankcase measurements, respectively. The sampling flow rates were selected to give the most rapid instrument response without causing air dilution of the samples. The pump was located after the sample cell. This arrangement placed the sampling system under a slight vacuum (about 20 mm of Hg with a flow rate of 0.3 cfm) so that airtight connections were required to prevent air leakage into the system. The sampling system was checked for leaks before each test.

The sampling system upstream from the analyzer was ¼-in. copper tubing. Preliminary tests indicated that there was no detectable difference in measured hydrocarbon concentrations using either stainless steel or copper sampling systems.

Continuous Hydrocarbon Analysis — Exhaust and crankcase hydrocarbon concentrations were measured with a Beckman Model 15A nondispersive infrared analyzer equipped with a n-hexane detector and quartz optics.

A 13½-in. sample cell was used for exhaust hydrocarbon measurement. To eliminate any response from the carbon dioxide in exhaust gas, the analyzer was equipped with a 1½-in. interference cell, which was filled daily with carbon dioxide. A correction was made for the response of the n-hexane detector to water vapor in the exhaust gases. The temperature of the exhaust gas entering the analyzer was measured; the correction was determined by assuming the exhaust to be saturated at the entering temperature. From a previous water response calibration, the water correction applicable to the exhaust gas measurements reported in this paper was determined to be 20 ppm n-hexane for the detector used.

Sample and reference cells of 1¼-in. length were used

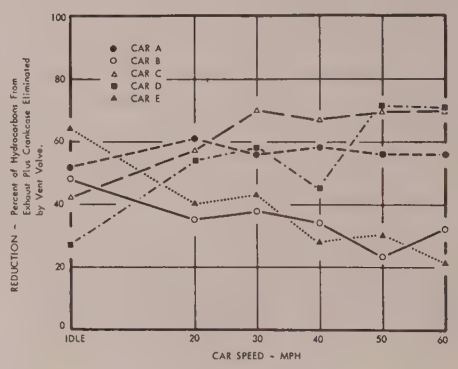


Fig. 18 — Reduction in hydrocarbon emission with internal crankcase ventilation

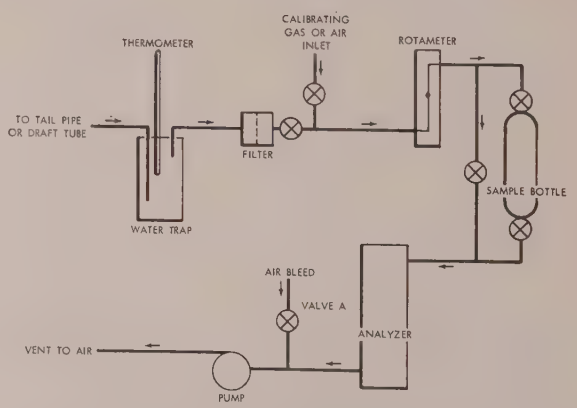


Fig. 19 — Sampling system

for measuring crankcase vent gas hydrocarbon concentrations. This cell length was chosen in order to have a linear calibration curve in the concentration range of 0–10,000 ppm n-hexane. Since no carbon dioxide interference cell was used when sampling crankcase gases, the n-hexane detector had not only water vapor response but carbon dioxide response. However, no carbon dioxide and water vapor corrections were applied, since the total corrections were estimated to be less than 1% of the crankcase hydrocarbon concentrations.

The infrared analyzer was calibrated daily with known concentrations of n-hexane in prepurified nitrogen. The calibrating gases obtained from The Matheson Co., Inc., were analyzed by gas chromatography. The calibrating gas concentrations for exhaust hydrocarbon measurements were 49, 100, 330, 504, and 1050 ppm n-hexane. Gases of concentrations of 1050, 2510, 5060, and 8750 ppm n-hexane were used for crankcase hydrocarbon calibration. The instrument gain was adjusted to obtain full-scale deflection on an Esterline-Angus recorder at 1050 and 8750 ppm n-hexane for exhaust and crankcase hydrocarbon measurements, respectively. The calibrating gas was swept through the analyzer at a low flow rate (0.01–0.03 cfm) until a steady reading on the recorder was obtained. In addition to these daily calibrations, calibration checks were made with either 1050 or 8750 ppm n-hexane before and after each run. If any variation occurred the average value was used. The maximum variation was less than 3%.

The response of the Beckman analyzer was found to depend upon the power source. To avoid this difficulty, calibrations were made with the power source used for exhaust and crankcase hydrocarbon measurements, an ATR converter and a 12-volt battery.

Gas Chromatographic Hydrocarbon Analysis — Individual hydrocarbon concentrations were determined by gas chromatographic analyses of grab samples of blowby gases. The samples were collected through a ¼-in. copper line from the road draft tube. The sample bottle was placed in series ahead of the Beckman analyzer and was purged with blowby gases until a steady analyzer deflection was obtained. Consequently, the gas chromatographic and Beckman determinations were made on equivalent samples. The grab sample was analyzed with a modification of the chromatographic apparatus described by Heaton and Wentworth[9] using four separate columns. The chromatographic column data and operating conditions are described in Table 4. The columns used and the hydrocarbons separated in each are:

1. Silica Gel-Alumina column: This column separates the C_1 to C_3 compounds plus the C_4 paraffins.
2. Phthalate column: This column separates the C_4 olefins and C_5 and C_6 compounds.
3. Silicone column: This column separates the C_5 to C_8 compounds.
4. Aroclor column: This column gives limited information on the heavier aromatics (to the trimethylbenzenes).

Typical chromatograms from the various columns for samples of blowby gases and fuel are presented in Figs. 2–5. On each chromatogram the compounds in each peak are identified. The peaks were identified by comparing the retention times of the various peaks with the retention times obtained with known compounds. Concentrations were calculated by comparing peak areas with the areas obtained for known concentrations.

Individual hydrocarbon concentrations and the total hydrocarbon concentration are listed in Table 5. These results were obtained from the chromatograms shown in Figs. 2–5. Besides the total concentration as ppm hydrocarbon (the sum of the individual concentrations) the concentration as ppm C_6 is listed. The sum of the individual concentrations times the average carbon number of the hydrocarbon mixture divided by 6 is equal to the concentration as ppm C_6. The concentration as C_8 is needed to compare Beckman-determined concentrations, which are reported as ppm n-hexane, and gas chromatographically determined concentrations.

The data in Table 5 show low concentrations of the C_1 to C_3 compounds and the extremely high concentrations of the fuel range hydrocarbons (C_4 to C_9 compounds). In other words, crankcase vent gas hydrocarbons are primarily those in the fuel.

Chromatograms of the fuel used are presented in Figs. 2–5. These chromatograms were obtained by injecting a small amount of the fuel with a syringe into the carrier-gas stream ahead of each column. No attempt was made to inject equal amounts of fuel into each column.

A number of crankcase vent gas samples were analyzed by gas chromatography to convert the Beckman-measured hydrocarbon concentrations to absolute hydrocarbon concentrations.

Continuous Absolute Hydrocarbon Analysis — The quantity actually measured with a nondispersive infrared analyzer is the amount of infrared energy in the transmitted spectral regions absorbed by the absorption bands of the sample gas that overlap the absorption bands of the gas in the detector. The infrared absorption bands utilized by n-hexane-filled detectors with quartz optics result from C—H stretching vibrations. Although all hydrocarbons have these absorptions bands, the position and intensity of each band depends upon the molecular structure. In the instrument used, different hydrocarbons have different relative responses on a C—H bond or carbon basis. The degree of similarity in molecular structure of a particular hydrocarbon and n-hexane determines the amount of band overlap and hence the response on a carbon basis relative to n-hexane. The response of the n-hexane detector to n-hexane is arbitrarily defined as 100. As shown in Table 6, it has been determined that the paraffinic hydrocarbons,

Table 4 — Chromatographic Columns, Packing and Operating Conditions

	Silica Gel-Alumina	Phthalate	Silicone	Aroclor
Packing — Solid	50% (wt) Silica gel (Davison No. 950) 50% (wt) Alumina (Alcoa No. F-20)	"Columpak" (Fisher)	"Columpak"	"Columpak"
— Liquid	None	50% (wt) di-isooctylphthalate 50% (wt) di-n-decylphthalate	Silicone No. 550 (Dow Corning)	Aroclor No. 1248 (Monsanto)
— Wt % Liquid	—	33	14	14
Length, ft[a]	5	20	10	5
Operating Temperature F	75	75	75	150
Flow Rate,[b] ml/min (STP)	150	150	75	75
Sample Volume, ml (STP)	27	27	56	56

[a] All Columns were ¼-in. copper tubing.
[b] Oxygen carrier gas.

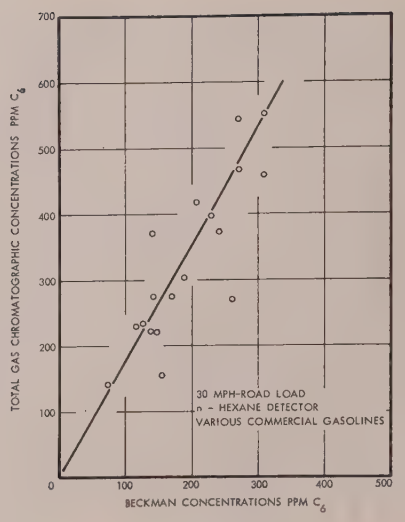


Fig. 20—Comparison of chromatographic and Beckman analyses of exhaust gases

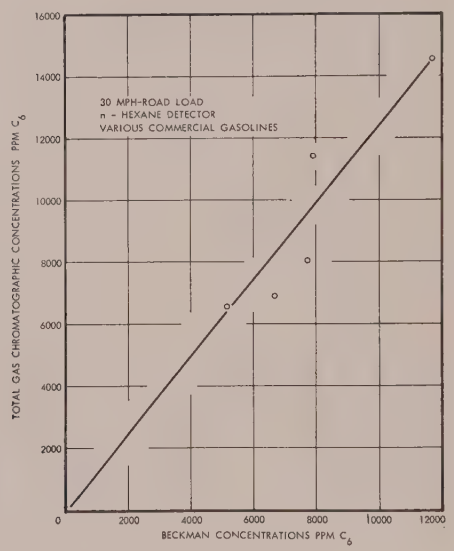


Fig. 21—Comparison of chromatographic and Beckman analyses of crankcase vent gases

Table 5 — Typical Blowby Hydrocarbon Analysis
(Engine of Make Z — Dynamometer Test)

Components	Hydrocarbon Concentration, ppm HC
Methane	26
Ethane	8
Ethylene	57
Propane	—
Acetylene	Trace
Propylene	90
i-Butane	90
Methylacetylene, Propadiene	—
n-Butane, i-Butene	612
2-Butene (trans), 1,3-Butadiene	52
2-Butene (cis)	65
i-Pentane, 3-Methyl-1-butene	1620
n-Pentane, 1-Pentene	606
2-Methyl-1-butene	200
2-Pentene, 2,2-Dimethyl-3-butene	216
2-Methyl-2-butene, 2,2-Dimethylbutane, 2-Methyl-1,3-butadiene	604
2-Methylpentane, 2,3-Dimethylbutane, 2-Methyl-4-pentene	2080
Cyclopentane, 2,3-Dimethyl-1-butene, 2-Methyl-3-pentene	—
3-Methylpentane	1190
n-Hexane, 1-Hexene, 2-Methyl-1-pentene	1660
3-Ethyl-1-butene, 2-Hexene, 3-Hexene, 2,2-Dimethylpentane, 2-Methyl-2-pentene, 2,4-Dimethylpentane	386
Methylcyclopentane, 3-Methyl-2-pentene, 2,3,3-Trimethylbutane, 2,3,3-Trimethyl-1-butene, 2,3-Dimethyl-2-butene, 3,4-Dimethyl-1-pentene	360
3,3-Dimethylpentane	—
2,3-Dimethylpentane, 5-Methyl-1-hexene, 3,4-Dimethyl-2-pentene, Cyclohexane, 3-Methylhexane	430
2,4,4-Trimethylpentane, 3-Ethylpentane, 5-Methyl-2-hexene	580
n-Heptane, 1-Heptene, Cyclohexene	429
Benzene	238
2,4,4-Trimethyl-1-pentene, 3-Heptene	—
Methylcyclohexane, 2,4,4-Trimethyl-2-pentene, 2,2,3-Trimethylpentane	390
2,3,4-Trimethylpentane	725
3-Methylheptane, 2,2,5-Trimethylhexane	310
n-Octane	—
1-Methylcyclohexene, 2-Ethyl-1-hexene, 1-Octene	—
Toluene, 2,3-Dimethyl-2-hexene, 2,6-Dimethyl-3-heptene, 2-Octene (cis)	472
2-Octene (trans)	—
Ethylbenzene	—
m-Xylene, p-Xylene, o-Xylene, i-Propylbenzene	250
Total Hydrocarbon Concentration ppm HC	13,746
ppm C_6	14,689
Average Carbon Number	6.4

with the exception of methane, have responses on a carbon basis nearly equal to the response of n-hexane. The low molecular weight olefins and aromatics, on the other hand, have very low responses relative to n-hexane. Long chain olefins and large side chain aromatics have somewhat higher relative responses due to the introduction of paraffinic C—H bonds into the structure. Similar findings have been reported in the literature.[11]

Since the response of the Beckman analyzer equipped with a n-hexane detector depends on hydrocarbon type as well as concentration, the composition of the sample affects the indicated hydrocarbon concentrations. Therefore, the indicated hydrocarbon concentration is an abitrary number. This number is somewhat larger than the paraffinic hydrocarbon content of the sample and is certainly smaller than the total hydrocarbon content.

This arbitrary number can be converted to an absolute concentration by using a correlation between simultaneously determined Beckman measurements and gas chromatographic total concentrations. The gas chromatographic technique has equal responses to all hydrocarbons; therefore, an absolute concentration can be established.

Correlations between the arbitrary Beckman concentrations and the absolute total gas chromatographic concentrations of exhaust and crankcase gas samples are shown in Figs. 20 and 21. In both of these figures the results are limited to comparisons of analyses made at 30 mph-road load test conditions. The correlations for exhaust hydrocarbon concentrations is shown in Fig. 20. Both the Beckman and gas chromatographic concentrations are plotted as ppm C_6. From the slope of the line it can be determined that the Beckman measurements should be multiplied by 1.8 to convert to absolute concentrations.

The crankcase vent hydrocarbon correlation is shown in Fig. 21. Simultaneous Beckman and gas chromatographic determinations were made on samples from Cars A, B, C, D, and E, as well as an engine of Make Z. The test points lie near a straight line whose slope is 1.25. Therefore, the arbitrary Beckman crankcase vent hydrocarbon determinations are converted to absolute concentrations when multiplied by 1.25.

Exhaust Flow Rate Measurement — Exhaust flow rates were calculated from the airflow rates by accounting for the change in volume due to the change in molecular composition going through the engine, and for that portion of the intake air in the engine blowby.

Airflow was calculated from fuel flow and air/fuel ratio

[11] "Application of Continuous Infrared Instruments to Analysis of Exhaust Gas," by B. W. Sturgis, W. F. Biller, J. W. Bozek, and S. B. Smith. Paper presented at SAE Annual Meeting, Detroit, January, 1958.

Table 6 — Relative Responses of n-Hexane Detectors
(Determined at Concentrations Less Than 1000 ppm C)

Methane	27	1-Pentene	56
Ethane	98	2-Pentene	60
Ethylene	8	n-Hexane	100
Acetylene	2	n-Heptane	107
Propane	116	1-Heptene	76
Propylene	37	Benzene	3
n-Butane	115	Toluene	17
i-Pentane	100	p-Xylene	30
n-Pentane	106		

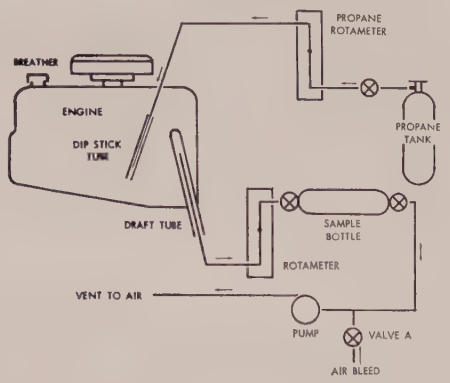


Fig. 22 — Tracer gas system for measuring ventilation flow rates

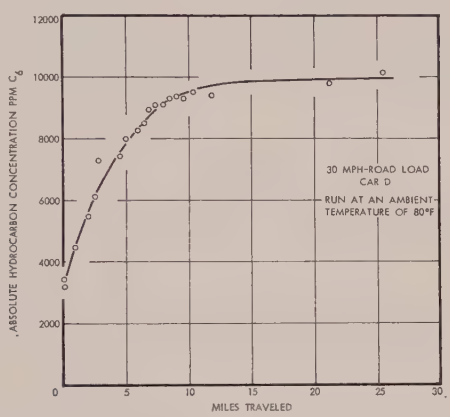


Fig. 23 — Effect of car warmup on blowby hydrocarbon concentration

measurements at each test condition. The fuel flow was measured with a volumetric device over a level course for ½ or 1 mile, depending upon the test speed. Grab samples of the exhaust were collected simultaneously from the tail pipe. These samples were analyzed by Orsat techniques to determine air/fuel ratios.

Crankcase Ventilation Rate Measurement — Crankcase ventilation rates were determined by a tracer gas technique. Chromatographic analyses have shown that engine exhaust and crankcase vent gases contain less than 1 ppm propane. Furthermore, chromatographic techniques for quantitative analysis of propane in concentrations less than 1000 ppm were readily available. Accordingly, propane was used as the tracer gas. A schematic diagram of the apparatus used is shown in Fig. 22.

The propane was metered at a constant flow rate through a 3/16-in. copper line into the crankcase through the dip stick tube. The propane rotameter was calibrated with a soap bubble flow meter. A grab sample of crankcase gases was collected in a clean 250-ml glass bottle at a 0.03-cfm flow rate through the ¼-in. copper sample line inserted in the road draft tube. This is the same sample line used in measuring crankcase hydrocarbon concentrations. The grab sample was analyzed chromatographically for propane. From the propane concentration in the draft tube sample and the propane input flow rate, the crankcase vent flow rate was calculated.

The merit for this system of flow rate measurement is the small disturbance imposed on the conventional crankcase ventilation system. The area of the sample tube was less than 4% of the area of the draft tube. The sampling rate was less than 5% of the crankcase ventilation rate, and the rate of propane addition was less than 0.5% of the crankcase ventilation rate. With these small alterations, it would be expected that the flow measuring technique would give flow rates representative of normal ventilating conditions.

In most cases the cars were tested with normal ventilation. However, when there was a possibility of backflow out of the breather, the breather was plugged. Under these conditions, flow rates and hydrocarbon concentrations measured were those of blowby gases rather than those of crankcase vent gases. Possible backflow conditions existed at idle and 20 mph-road load for all the test cars. Consequently, the breathers were plugged at these conditions. On Cars D and E, because of the mechanical design of the breather system, difficulties were encountered in detecting backflow. For these cars, the breathers were plugged for all tests. The breathers were not plugged for any tests involving internal crankcase ventilation systems.

Road Test Procedures — Three separate runs were made with each vehicle because of the different instrumentation and equipment needed for the various measurements.

One run was made to determine crankcase ventilation rates. The propane cylinder, rotameter, lines, sample bottle, and pump were installed. During a warmup of about 10 miles, propane was added to the crankcase at a constant rate. After warmup the vehicle was held at a constant speed for at least 2 min. During the last minute the sample was collected and subsequently analyzed for propane. Measurements were made at the following operating conditions: idle; 20, 30, 40, 50, and 60 mph-road load; and 30 mph-10 in. of Hg manifold vacuum and full throttle.

Another run was made to measure the hydrocarbon concentration in the crankcase vent or blowby gases after warmup. The results of tests to determine the effect of car warmup are shown in Fig. 23. All measurements were made at 30 mph-road load. The blowby hydrocarbon concentrations increased rapidly with increasing mileage and an equilibrium concentration was reached at about 10 miles. Consequently, the variation in blowby hydrocarbon concentrations during warmup must be considered in making concentration measurements. In the present work the test cars were warmed up for 10 miles prior to making any measurements. With the short sample and reference cells in the Beckman analyzer, the gases were sampled at the operating conditions listed in the preceding paragraph until a steady hydrocarbon concentration resulted.

During the final run, grab samples of exhaust gases for Orsat analyses were collected, and exhaust hydrocarbon concentrations and fuel flow rates were measured. The Beckman analyzer with the long sample cells was used to measure exhaust hydrocarbons. The operating conditions were the same as listed previously.

APPENDIX II

Blowby Process

A consideration of the engine blowby process appears to offer reasons for the high hydrocarbon content of engine blowby gases. The pressure differential across the pistons of the automotive 4-stroke internal-combustion engine indicates that most of the blowby must occur during the compression and expansion strokes. At top center the direction of piston travel changes, and at about this time piston side thrust reversal occurs. Both of these factors would be expected to interfere with the piston-ring sealing near the time of maximum cylinder pressure (shortly after top dead center). Accordingly, one would expect maximum blowby rates to occur near top center. However, factors such as

ring flutter, engine design, and operating conditions may vary the time of occurrence of maximum blowby rates.

Flame photographs of the combustion process have shown that even at full throttle the charge in the combustion chamber does not burn "completely" until after the piston has reached top dead center.[12,13] Because of the somewhat central location of the spark plug and the generally spherical nature of the propagation of the flame front, the last portion of the charge to burn is located in the angular section of the combustion chamber bounded by the piston top and cylinder wall. Thus, a consideration of flame travel and cylinder pressures offers an explanation why blowby gas consists largely of unburned charge. This explanation was first proposed by Cutter[7] in 1933 and more recently by Spindt, Wolfe, and Stevens.[8]

An additional factor influencing the hydrocarbon content of blowby gases may be that of wall quenching of the flame. Calculations of probable flame quenching distances by Agnew[14] based on the work of Friedman and Johnson[15] indicate that at engine cylinder pressures typical of road load operation, the flame may not propagate between metal plates some 0.020 in. apart. These calculations are consistent with direct observations of flame quenching distances in engines.[16] Since the engine used in obtaining the data of Tables 1 and 2 has a 0.015-in. land clearance (Fig. 24), the flame may never enter that portion of the combustion chamber bounded by the piston, the top ring, and the cylinder wall. During the compression and expansion strokes, the piston ring is normally held in contact with the lower side of the ring groove and the cylinder wall by gas pressure. Thus, the total quench volume extends into the ring groove to include all of volume A (Fig. 24). For this engine, volume A is approximately 0.11 cu in.

Engine blowby rate measurements corresponding to the data of Tables 1 and 2 have shown that the volume of the blowby gas at atmospheric pressure was about 0.14 cu in. for each complete cycle of each cylinder. Accordingly, allowing for the increase in gas density in the combustion chamber and assuming uniform blowby around the ring perimeter, it is clear that flame quenching between the piston, the ring, and the cylinder wall could also account for the high hydrocarbon content of engine blowby gases.

An additional factor that may be of importance is the flow of gases through the piston ring gap.

The above discussion has presented a number of factors which must be considered in arriving at an explanation for the composition of engine blowby gases. However, either flame travel or wall quenching phenomena could account for the observations that blowby gases are largely carbureted mixture.

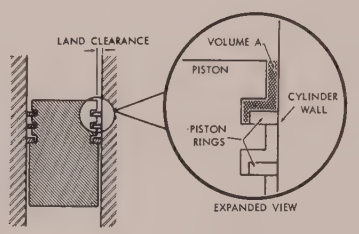


Fig. 24 — Detail of piston, piston ring, and cylinder wall

[12] "Flame Photographs of Autoignition Induced by Combustion Chamber Deposits," by L. L. Withrow and F. W. Bowditch. SAE *Quarterly* Transactions, Vol. 6, October, 1952, pp. 724–752.

[13] "Some Aspects of Particle Ignition," by F. W. Bowditch, R. E. Wilson, and G. J. Nebel. Paper presented at API Division of Refining meeting, Houston, May, 1954.

[14] "Possibility of Wall Quenching as Cause for Unburned Hydrocarbons in Engine Exhaust," by W. G. Agnew. GM Research Laboratory Report, 1954.

[15] "Wall-Quenching of Laminar Propane Flame as Function of Pressure, Temperature, and Air-Fuel Ratio," by R. Friedman and W. C. Johnston. *Journal Applied Physics*, Vol. 21, 1950, p. 791–795.

[16] "Flame Quenching at Walls of Internal Combustion Engine," by W. A. Daniel. Sixth Symposium (International) on Combustion. Pub. by Reinhold Publishing Corp., New York, 1957.

APPENDIX III

Table 7 — Summary of Concentration, Flow Rate, and Emission Data on Car A
(1959 Make V — 7000 Miles)

	Standard			Internal Crankcase Ventilation		
Exhaust Gases	**Absolute HC Conc., ppm C_6**	**Flow, cfm**	**Absolute Emission lb C_6/hr**	**Absolute HC Conc., ppm C_6**	**Flow, cfm**	**Absolute Emission lb C_6/hr**
Idle	1230	8	0.12	760	10	0.09
20 mph road load	520	17	0.11	400	18	0.09
30 mph road load	450	26	0.14	380	27	0.13
40 mph road load	420	38	0.20	320	38	0.15
50 mph road load	350	51	0.22	280	51	0.18
60 mph road load	290	69	0.25	220	72	0.20
30 mph — 10 in. of Hg	320	63	0.25	240	61	0.18
30 mph — full throttle	320	111	0.43	270	113	0.38

	Standard			Internal Crankcase Ventilation		
Crankcase Gases	**Absolute HC Conc., ppm C_6**	**Flow, cfm**	**Absolute Emission lb C_6/hr**		**Flow, cfm**	**Absolute Emission lb C_6/hr**
Idle	8340[a]	0.7[a]	0.07[a]		0	0
20 mph road load	8750[a]	1.1[a]	0.12[a]			
30 mph road load	8250	1.4	0.14			
40 mph road load	7440	1.7	0.16			
50 mph road load	7190	2.0	0.18		0	0
60 mph road load	7500	2.3	0.21		0	0
30 mph — 10 in. of Hg	8120	1.9	0.19			
30 mph — full throttle	9120	2.5	0.28		0	0

[a] Breather plugged.

Table 8 — Summary of Concentration, Flow Rate, and Emission Data on Car B
(1959 Make W — 10,000 Miles)

	Standard			Internal Crankcase Ventilation		
Exhaust Gases	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr
Idle	650	13	0.10	580	13	0.09
20 mph road load	580	21	0.15	580	21	0.15
30 mph road load	470	27	0.16	500	26	0.16
40 mph road load	450	41	0.23	450	41	0.23
50 mph road load	420	52	0.27	440	56	0.31
60 mph road load	400	75	0.37	420	76	0.40
30 mph — 10 in. of Hg	400	67	0.33	370	63	0.29
30 mph — full throttle	400	102	0.50	420	97	0.51

	Standard			Internal Crankcase Ventilation		
Crankcase Gases	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr		Flow, cfm	Absolute Emission, lb C_6/hr
Idle	8210[a]	1.0[a]	0.08[a]		0	0
20 mph road load	8000[a]	1.1[a]	0.08[a]			
30 mph road load	6340	1.7	0.10			
40 mph road load	6090	2.1	0.12			
50 mph road load	5880	2.3	0.13			
60 mph road load	6190	2.5	0.14		0	0
30 mph — 10 in. of Hg	6900	2.2	0.14			
30 mph — full throttle	8210	2.6	0.20		0	0

[a] Breather plugged.

Table 9 — Summary of Concentration, Flow Rate, and Emission Data on Car C
(1959 Make X — 12,000 Miles)

	Standard			Internal Crankcase Ventilation		
Exhaust Gases	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr
Idle	1250	8	0.12	1260	11	0.17
20 mph road load	500	18	0.11	440–610	19	0.10–0.14
30 mph road load	310	27	0.10	270	27	0.09
40 mph road load	280	41	0.14	250	40	0.12
50 mph road load	300	60	0.22	250	54	0.17
60 mph road load	250	74	0.23	220	71	0.19
30 mph — 10 in. of Hg	290	67	0.24	250	66	0.20
30 mph — full throttle	270	103	0.34	260	97	0.31

	Standard			Internal Crankcase Ventilation		
Crankcase Gases	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr		Flow, cfm	Absolute Emission, lb C_6/hr
Idle	10,000[a]	1.4[a]	0.17[a]			
20 mph road load	10,190[a]	1.4[a]	0.17[a]			
30 mph road load	9,560	1.6	0.19			
40 mph road load	8.750	2.2	0.24			
50 mph road load	8,750	3.1	0.33			
60 mph road load	8,380	3.9	0.40		0	0
30 mph — 10 in. of Hg	11,120	2.7	0.37		0	0
30 mph — full throttle	11,440	3.3	0.46		(Backflow Out Breather)	

[a] Breather plugged.

Table 10 — Summary of Concentration, Flow Rate, and Emission Data on Car D
(1950 Make Y — 59,000 Miles)

	Standard			Internal Crankcase Ventilation		
Exhaust Gases	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr	Absolute HC Conc., ppm C_6	Flow, cfm	Absolute Emission, lb C_6/hr
Idle	950	6	0.07	990–1390	6	0.07–0.10
20 mph road load	560	13	0.09	370	14	0.06
30 mph road load	340–580	20	0.08–0.14	130–590	20	0.03–0.15
40 mph road load	220	33	0.09	180	33	0.07
50 mph road load	170	46	0.10	100	49	0.06
60 mph road load	120	62	0.09	80	66	0.07
30 mph — 10 in. of Hg	210	43	0.11	140	43	0.08
30 mph — full throttle	320	69	0.27	290	68	0.24

cont. on following page

Table 10, cont.

Crankcase Gases	Standard: Absolute HC Conc., ppm C_6	Standard: Flow, cfm	Standard: Absolute Emission, lb C_6/hr
	(Measured with breather plugged)		
Idle	10,250	0.4	0.05
20 mph road load	10,000	0.4	0.05
30 mph road load	10,000	0.5	0.06
40 mph road load	10,620	0.6	0.08
50 mph road load	11,250	0.8	0.11
60 mph road load	10,940	1.0	0.13
30 mph — 10 in. of Hg	9,880	0.6	0.07
30 mph — full throttle	11,880	0.9	0.13

Table 11 — Summary of Concentration, Flow Rate, and Emission Data on Car E
(1955 Make Y — 41,000 Miles)

Exhaust Gases	Standard: Absolute HC Conc., ppm C_6	Standard: Flow, cfm	Standard: Absolute Emission, lb C_6/hr	Internal Crankcase Ventilation: Absolute HC Conc., ppm C_6	Internal Crankcase Ventilation: Flow, cfm	Internal Crankcase Ventilation: Absolute Emission, lb C_6/hr
Idle	500	5	0.03	710	6	0.05
20 mph road load	410	20	0.10	380	23	0.11
30 mph road load	280	31	0.11	290	30	0.11
40 mph road load	200	43	0.10	250	44	0.14
50 mph road load	180	58	0.13	200	61	0.15
60 mph road load	160	75	0.15	200	78	0.19
30 mph — in. of Hg	220	51	0.14	330	51	0.21
30 mph — full throttle	230	73	0.21	390	76	0.36

Crankcase Gases	Standard: Absolute HC Conc., ppm C_6	Standard: Flow, cfm	Standard: Absolute Emission, lb C_6/hr	Internal Crankcase Ventilation: Absolute HC Conc.,[a] ppm C_6	Internal Crankcase Ventilation: Flow,[b] cfm	Internal Crankcase Ventilation: Absolute Emission, lb C_6/hr
	(Measured with breather plugged)				(Backflow out breather)	
Idle	6750	1.3	0.11	—	0	0
20 mph road load	6000	2.0	0.15	6000	0.5	0.04
30 mph road load	7060	2.2	0.19	7060	0.7	0.06
40 mph road load	6690	2.3	0.19	6690	0.8	0.07
50 mph road load	7190	2.7	0.24	7190	1.2	0.11
60 mph road load	7500	3.4	0.32	7500	1.9	0.18
30 mph — in. of Hg	7190	3.4	0.30	7190	0.9	0.08
30 mph — full throttle	7500	3.5	0.32	7500	0.8	0.07

[a] Estimated to be the same concentration as blowby gases.
[b] Calculated from blowby rates and vent valve flow rates.

DISCUSSION

Numerous Research Programs Are Underway on Air Pollution

— C. M. Heinen
Chrysler Corp.

WHEN THE RESULTS of these investigations of crankcase emissions were made available to the Vehicle Combustion Products Committee of the AMA, the industry reaction was given most eloquently. In a period of less than six months the industry committee reviewed the data, made confirmatory investigations, examined the engineering difficulties, and recommended that action be taken by releasing crankcase breather control systems for use on automobiles offered for sale in California. As has been publicly announced, the Engineering Advisory Committee and the Board of Directors of the AMA acted on this recommendation and it is expected that all 1961 model American-made automobiles offered for sale in California will be so equipped.

This action on the control of vehicle emissions does not, however, represent the whole program of the automotive industry.

The problem of measurement of low concentrations of polluting materials in vehicle emissions and the problems of analysis for organics and other components are still subjects for intensive cooperative effort in the industry. We do not regard analytical techniques as being sufficiently accurate under all conditions and there is such variability from vehicle to vehicle and from one operating condition to another that many tests must be run on each variable before even approximate values can be determined. Incidentally, this is what accounts for the generous use of such terms as "substantial," "approximately," "of the order of," and the like, in reports on the vehicle emission problem.

Through the CRC the industry representatives are engaged in an active study of the nondispersive infrared instruments generally used in our work on hydrocarbon, carbon monoxide, and carbon dioxide emissions.

Currently underway is a program to establish a prototype inspection station in Los Angeles in cooperation with local authorities which should help to improve instrumentation and measuring techniques, as well as provide additional data on emissions and develop information on how

periodic inspection of vehicles might be conducted successfully.

Other major aspects of the current program in the industry include a project of our Induction System Task Group which is studying evaporation losses through carburetor vents and other parts of the fuel system. The major effort, of course, is directed at the development of satisfactory exhaust control devices. In 1959 SAE papers were presented indicating the technical feasibility of using high-temperature catalysts, low-temperature catalysts, and direct flame-type afterburners. During the past year all three of these systems have been subjected to intensive engineering development aimed at taking them from the stage of technical feasibility to the stage of commercial practicability.

In connection with these efforts the industry has continued to improve its methods of testing and evaluating such devices in traffic cycles, which include warmup and cooling periods as well as other conditions. The committee has worked also at the development of a suitable screening procedure or basis for evaluation of new catalysts which might be applied in the exhaust cleanup effort.

There are many other aspects to the industry committee activities which have been continued on an intensified basis during the past year. For instance, the Vehicle Combustion Products Committee has supplied technical cooperation to the California State Department of Health in its work on developing exhaust emissions standards. We have continued to work with the Federal Government on engineering aspects of its air pollution health studies as they relate to motor vehicles.

In summary, the industry regards the crankcase breather control device as a very substantial step in the reduction of objectionable emissions from motor vehicles but will continue its activities undiminished in the other areas of research and engineering development.

Blowby Gases Contribute To Smog Potential

— Leslie A. Chambers
Los Angeles County
Air Pollution Control District

THE AUTHORS have brought to light the importance of blowby emissions in the automotive smog problem. While this discovery, as the paper indicates, was somewhat by chance, it still was most opportune and will mark an important date in the campaign to control smog from motor vehicles.

The Air Pollution Control District findings, on a very limited number of vehicles, agree in general with the conclusions of the paper. The preliminary District data indicate that blowby losses are approximately one-third of the total automotive losses excluding evaporation. This paper estimated "that the internal crankcase ventilating system eliminates 40% of the engine hydrocarbon emissions exclusive of carburetor vent losses." There were no data given in the paper as to how the individual findings were weighed to arrive at this 40% estimate, but it is our feeling that a conservative weighing was used in view of some extremely high blowby volumes recorded in the paper. The Air Pollution Control District has not observed blowby volumes to be as high as those reported by the authors. Data in Tables A and B were obtained with a Sprague Gas Meter and confirmed by an orifice metering system. The hydrocarbons were recorded with a Liston-Becker, Model 15, nondispersive infrared analyzer. Since it was difficult to correlate the above steady-state conditions with those that actually occur on the road, a vehicle was equipped for determining blowby under actual road conditions. Two methods were used.

The first method was to insert a Sprague meter in the vent of a closed crankcase system. During deceleration (whenever the manifold vacuum was above that of idle and the throttle closed), a solenoid valve vented the crankcase to atmosphere. Thus, during deceleration periods, air would be allowed to be drawn into the crankcase. The blowby gases were sampled at a continuous rate into an evacuated cylinder (2 l/min). Table C shows data collected by this method.

From this method, with a very limited number of runs, it appears that blowby losses are approximately 50% of that coming from the exhaust pipe—or one-third of the total hydrocarbon losses, exclusive of evaporation. It must be understood that the above determinations have been based on only one vehicle. Since this vehicle is extremely

Table A — Engine A 1957 6-Cyl with Approximately 10,000 Miles

Phase	Manifold Vacuum, in. of Hg	Rpm	Blowby Gases	
			cfm	HC, ppm
Idle	19.0	500	0.15	7800
Cruise, 25	18.0	1100	0.33	7800
Cruise, 35	15.0	1500	0.48	8200
Cruise, 45	16.0	1900	0.44	8500
Cruise, 55	15.0	2300	0.60	8500
Acceleration	0.2	1200	1.24	8300
Acceleration	0.4	1600	1.24	8300
Acceleration	0.6	2000	1.22	8600
Deceleration	25.0	1500	0.23[a]	—

[a] Air was drawn into the crankcase during the deceleration operation. Hence, hydrocarbon emission was considered none.

Table B — Engine B 1957 V-8 with Approximately 5000 Miles

Phase	Manifold Vacuum, in. of Hg	Rpm	Blowby Gases	
			cfm	HC, ppm
Idle	19.0	500	0.26	6500
Cruise, 25	18.0	1100	0.77	6200
Cruise, 35	15.0	1500	1.09	6200
Cruise, 45	16.0	1900	0.98	6300
Cruise, 55	15.0	2300	1.05	6600
Acceleration	0.2	1200	1.84	7800
Acceleration	0.4	1600	1.90	7800
Acceleration	1.1	2000	1.90	7800
Deceleration	25.4	1500	0.18[a]	—

[a] Air was drawn into the crankcase during the deceleration operation. Hence, hydrocarbon emission was considered none.

Table C — Data Collected from 6-Cyl 1956 Vehicle with Odometer Reading of 49,000 Miles

Run No.	Blowby Gases					
	cu ft/hr[a]	HC, ppm	CO_2 %	CO, %	g/hr (as hexane)	Blowby Gases / Exhaust Gases × 100 = %[b]
1	64.0	6000	3.2	0.9	38.4	$\frac{38.4}{82.0} \times 100 = 46.7$
2	63.2	6700	3.4	1.1	42.3	$\frac{42.3}{82.0} \times 100 = 51.5$
3	67.0	6730	3.2	1.2	45.0	$\frac{45.0}{82.0} \times 100 = 54.9$
Average	64.7	6476	3.2	1.0	41.9	51.0

[a] Sprague Gas Meter was used for the volume measurement.
[b] Exhaust Gases:
Volume = 1500 cu ft/hr at 75 F
Hydrocarbon = 550 ppm (as hexane)

$$(1500)\ (28.3)\ (0.000550)\ \left(\frac{86}{24.4}\right) = 82.0 \text{ g/hr}$$

Carbon dioxide = 12.2%
Carbon monoxide = 3.5%

low in average hydrocarbon exhaust emissions, the blowby emissions expressed as a percentage of the exhaust emissions would perhaps be higher than might occur from a vehicle with higher average hydrocarbon exhaust emissions.

The same 1956 6-cyl vehicle was again analyzed by an entirely different technique. A proportional sampling valve[a] was utilized to collect an exhaust sample in an evacuated container under similar road conditions both with and without diversion of the blowby into the exhaust stream. The proportioning valve sampled at a point approximately 30 in. from the end of the tailpipe, and the blowby stream was piped into the exhaust system immediately following the muffler. The vehicle was run on a route typical of Los Angeles County driving conditions. Data from this run are shown in Table D. It will be noted that the hydrocarbon value from the exhaust stream was an average of 460 ppm (hexane equivalent), while the value from the exhaust stream plus the blowby emissions was 673 ppm. The weight of hydrocarbons that these concentrations represent amount to approximately 50% of the exhaust emissions or one-third of the total hydrocarbons from the test vehicle, exclusive of the evaporation losses.

Since the two methods utilized for determining blowby values on the road gave equivalent results, we feel that either system can be relied upon for road blowby determinations. However, we prefer the use of the proportional sampling valve, as this eliminates some imperfections of the gas meter system in determining blowby concentrations.

Preliminary checks were made with a 50-l flask to determine the relative eye-irritating potential of blowby gases. Table E shows the time, in seconds, that individuals first noticed eye irritation. In general, the lower the number of seconds, the greater the eye irritation effect. A standard test fuel was utilized and the flask was treated with 1, 3, and 5 ppm of hydrocarbon. While the results must again be considered as preliminary in nature, they do indicate that the blowby eye irritating potential falls between that of the exhaust gas and the fuel itself. No tests have yet been conducted using the 1000-cu-ft environmental chamber.

Since the literature, as well as this paper, shows extreme variations in the amount of blowby from individual vehicles, it is essential that the blowby problem be specifically studied as it relates to the cars in Los Angeles County. It is, therefore, recommended that future work contemplated in the determination of the condition of Los Angeles County vehicles be extended to include blowby emissions.

Table D — Data Collected from 6-Cyl 1956 Vehicle with and without Blowby Gases Added to Exhaust Gases

Run No.	Exhaust Gases					Exhaust Gases + Blowby Gases				
	cu ft/hr	HC, ppm	CO_2, %	CO, %	g/hr (hexane)	cu ft/hr	HC, ppm	CO_2, %	CO, %	g/hr (hexane)
1	1500	430	12.3	2.5	64.2	1565	660	12.0	3.1	103
2	1500	520	10.9	3.0	77.7	1565	670	11.8	3.3	105
3	1500	440	10.9	3.4	65.6	1565	690	11.2	3.4	108
Av.	1500	460	11.4	3.0	69.1	1565	673	11.6	3.3	105

[a] Blowby Gases 65 cu ft/hr at 75 F
Exhaust Gases 1500 cu ft/hr at 75 F

$$\frac{(105-69)}{69} \times 100 = 52\%$$

(Blowby gases represent 52% of the exhaust gases)

Table E — Eye Irritation Data from Blowby Exhaust Gases

	Fuel 3	Blowby 3	Exhaust 3
1 ppm	121 sec	74 sec	54 sec
3 ppm	52 sec	45 sec	31 sec
5 ppm	29 sec	28 sec	23 sec

Since the burning of blowby gases in the engine met with disfavor some years ago because of carburetor gumming when introduced into the carburetor and intake valve, and combustion-chamber deposits when introduced into the manifold, it is assumed that the automotive industry will investigate these areas before the introduction of blowby control devices. Field experience with blowby recirculation will soon be available to the Air Pollution Control District as the Board of Supervisors has ordered installation of devices for this purpose on the entire Los Angeles County fleet.

The discovery of the blowby smog potential is certainly not "an error of omission," but is equally an "error of inclusion" as many European vehicles (including at least two Russian makes) are equipped with a means for burning blowby gases; and one very popular American truck manufacturer has been utilizing a means of disposing of blowby gases for many years without the realization that the system was beneficial from a smog control standpoint.

Crankcase Emission Control Another Step in Smog Control

— John A. Maga
California State Department of Public Health

THE AUTHORS' findings indicate that present estimates of hydrocarbon losses may need to be substantially changed. The paper also illustrates the complex problems that are inherent in the control of hydrocarbon emissions from motor vehicles. Not only are there many difficulties in measuring hydrocarbons, estimating total losses to the atmosphere, and evaluating the importance of the various hydrocarbons emitted, but the motor vehicle itself discharges pollutants from several points. The quantities discharged may vary widely from vehicle to vehicle and under different operating conditions.

It had been assumed by many investigators and California air pollution agencies that crankcase emissions represented a very small per cent of the total hydrocarbon losses from motor vehicles. This conclusion was based on estimated hydrocarbon concentration in the crankcase emissions much lower than that indicated by the authors. The paper presents data for five automobiles. It is not known if this data is representative of the motor vehicle population in California. However, there seems to be little question that blowby gases may be a greater source of hydrocarbons than had been believed and that these gases must be considered in any program to control hydrocarbon emissions from motor vehicles.

The control of pollutants from motor vehicles is an important part of California's attack on its air pollution problems. To date major emphasis has been placed on the hydrocarbons emitted from the tailpipe. Some attention also has been directed to fuel composition and evaporative losses from the fuel tank and carburetor. Hydrocarbon emissions from the crankcase do not appear to present as many control problems as do tailpipe emissions, and different types of afterburners from those now contemplated will not be required. A logical program for California to follow would include the feeding of crankcase gases back to the induction system of the engine, as proposed by the authors, together with the installation of effective afterburners for tailpipe emissions. It, of course, should also be feasible to pass the crankcase gases directly through an exhaust afterburner. This latter approach may be a less desirable one. Accepting the authors' estimate that crankcase hydrocarbon losses are approximately two-thirds those from the tailpipe, it should be possible to effect a slight improvement in fuel economy by salvaging the hydrocarbons of blowby gases.

The California State Department of Public Health re-

[a] "Proportional Sampling System for Collection of Integrated Auto Exhaust Gas Sample," by J. C. Chipman and M. T. Massey. Paper presented at Air Pollution Control Assoc. 52nd Annual Meeting, Los Angeles, June, 1959.

Table F — Blowby Flow Rates

Cal Research	Engine Miles	Blowby Flow Rate, cfm	
		Idle	20 mph Road Load
1955 Make A	74,000	0.3	0.6
1957 Make B	25,000	0.5	0.6
1958 Make C	20,000	0.6	0.9
1958 Make D	17,000	0.6	0.5
1959 Make E	12,000	0.6	0.8
1959 Make F	13,000	0.5	0.7
Cal Research 6-car average		0.5	0.7
General Motors 5-car average		1.0	1.2
Ratio, $\frac{\text{General Motors}}{\text{Cal Research}}$		2	1.7

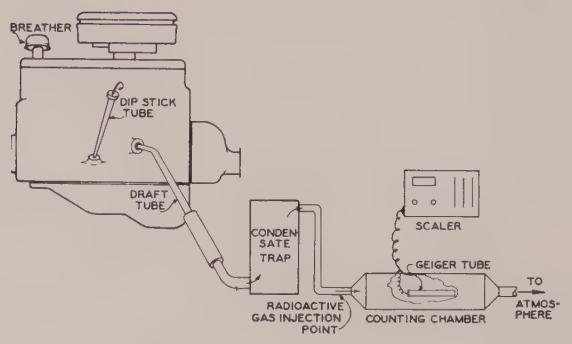


Fig. A — Measurement of blowby rate by total count method

cently adopted standards for emissions from motor vehicle exhausts, as required by the 1959 State Legislature. Standards for hydrocarbons were developed considering existing air quality, estimated present emission of hydrocarbons, and expected emissions in 1970. Crankcase gases were not included in estimates of hydrocarbon losses since it had been concluded they were only a small part of the total hydrocarbon losses. It is the policy of the Department that the standards will be refined, extended, or amended as new scientifically accepted information is available. During 1960 the motor vehicle standards will be re-evaluated in light of new data that will be available on air pollution, crankcase gases, and total vehicular emissions.

At first glance it would seem that if crankcase hydrocarbon emissions are as much as two-thirds of the amount of tailpipe emissions that the standards adopted by the State Department of Public Health should be changed materially. Actually, this may not be the case. In developing the standards, the Department considered that olefins were the important hydrocarbons to be controlled. Regulations of the Los Angeles County Air Pollution Control District will limit olefins in fuel to 12.5% by 1962. The exact role of fuel composition on exhaust olefins has not been established, but some studies have indicated that exhaust olefins are not directly related to those in the fuel. It is expected that the relationship of fuel olefins to crankcase emission olefins would be much more direct. The authors have shown that the hydrocarbons in blowby gases are very similar to those in the fuel. The control of fuel composition, therefore, should be effective in reducing olefins in the crankcase gases.

In setting the motor vehicle exhaust standards, the Department of Public Health assumed exhaust olefins were 45% of the hydrocarbons as measured by a hexane analyzer. If blowby gases are essentially mixtures of fuel and air as reported by the authors, the per cent olefins in blowby hydrocarbons should be about 12.5 in the Los Angeles area in 1962. The ratio of olefins in crankcase emissions to olefins in exhaust emissions would then be about one-fifth instead of the two-thirds for total hydrocarbons as given in the paper. Until 1962 olefin content of fuels, of course, will be greater than 12.5% in Los Angeles. Other California air pollution control agencies have not adopted regulations on olefins in fuel.

The proposed control of hydrocarbons from cranckcase emissions is an encouraging step in the solution of California's air pollution problem. The method suggested by the authors seems to be highly effective, practical, and inexpensive. Apart from consideration of the relative contributions of exhaust, blowby, and evaporative losses, the simplicity of the proposed blowby control device would appear to call for its early installation on California vehicles. Although we may argue how much benefit may be expected, it seems clear that definite reductions in hydrocarbons will result from the application of these findings to the air pollution problem in California.

It is important that studies of the nature reported in this paper be continued to answer many of the questions that remain on emissions from motor vehicles, their contributions to air pollution, and methods of emission control. As California's communities grow it will become increasingly difficult to maintain suitable air quality. Practical methods should be sought to eliminate the pollutants from all the sources on motor vehicles.

Blowby Gases Cause 20-25% Of Engine Hydrocarbon Emissions

— R. K. Stone
K. L. Kipp
J. Harkins
California Research Corp.

GENERAL MOTORS Research has conducted an excellent study of automotive crankcase vent emissions. It should be particularly gratifying to all concerned with control of air pollution that control of this source of unburned hydrocarbons appears controllable at relatively low cost. For this reason, their magnitude relative to other sources of automotive unburned hydrocarbon emissions is perhaps of only academic interest to many people. Nevertheless, we would like to make our comments on the magnitude of crankcase hydrocarbon emissions relative to exhaust gas emissions.

The authors have concluded that crankcase hydrocarbon emissions are approximatly 40% of the total engine emissions (exclusive of carburetor vent losses). This figure was obtained by measuring the flow rates and compositions of both the blowby and exhaust gases in five select cars. In reviewing the blowby data presented in the paper, it is apparent that their composition was roughly the same among the five cars. On the other hand, the flow rates varied as much as threefold at some operating conditions. We have recently completed tests measuring blowby rates at idle and 20 mph, road load, in six late model cars. These data were obtained with a radioactive tracer technique developed some years ago by Cal Research for measuring flow rates in liquids or gases.[b] Fig. A shows schematically how this test is conducted. The crankcase breather and dip stick tube were plugged for these tests.

It is of interest to look at the blowby data we obtained (Table F). The average General Motors data are included for comparison. First, it may be noted that we, too, find a variation in blowby rates; about twofold among the cars tested. However, the average blowby rates of our six cars are about one-half the averages for the five General Motors cars. If our figures for blowby rates are correct, then blowby hydrocarbon emissions may be on the order of 20–25% of the total from the engine, rather than 40% as found by General Motors.

The absolute exhaust hydrocarbon emissions in pounds

[b] "Engine Air Flow by Total-Count Method," by B. A. Fries, F. J. Davis, and D. E. Hull. Paper presented at SAE Annual Meeting, Detroit, January, 1959.

per hour were quite low for the five General Motors test cars as compared to the more generally accepted figures. A 300-car survey conducted by the Coordinating Research Council in Los Angeles in 1956 showed that 5–6% of the gasoline supplied was emitted from the exhaust unburned for an average driving cycle. This compares to an average of less than 3% calculated for the same driving cycle from the data on the five General Motors cars. One careful investigation of the exhaust hydrocarbon emissions found in the Los Angeles area, as reported in a 1959 SAE paper, ("Maintenance and Automobile Exhaust," by Fagley, et al.) found that the Los Angeles survey data were the right order of magnitude and attributed this high rate to the poor mechanical condition of the average automobile. The Los Angeles Air Pollution Control District and the California State Board of Public Health, after careful consideration of the available data, as well as recommendations of representatives of the petroleum and automotive industries, have also used a figure for exhaust hydrocarbon emissions which is consistent with the Los Angeles survey findings and almost twice the average value shown in the General Motors paper.

In our recent investigation of blowby rates we found that the magnitude of blowby hydrocarbon emissions is largely independent of the magnitude of exhaust hydrocarbon emissions. If the more accepted figure for exhaust hydrocarbon emissions is used together with the average GM Research Laboratories blowby emission data, we calculate that blowby gases contribute 20–25%, not 40% of the engine hydrocarbon emissions (exclusive of carburetor vent losses). Should average blowby rates be only one-half to two-thirds of those reported by General Motors, as our recent data show may indeed be the case, then blowby may actually contribute as little as 10–15% of the total engine hydrocarbon emissions.

Additional tests on a much larger car sample should be run to establish a more reliable estimate of the contribution of crankcase fumes to air pollution. Even though they may prove to be only 10–15% of the total engine hydrocarbons, as compared to the 40% suggested by General Motors, elimination of this source of air pollution may well be justified at this time.

Blowby Is Source Of Atmospheric Olefins

— W. L. Faith
Air Pollution Foundation

THE AUTHORS have called attention to a long-overlooked source of air pollution from motor vehicles. We are thus faced with the problem of the importance of these emissions in a community. I shall confine my comments largely to appraisal of this new information as it applies to the situation in California, especially in the Los Angeles area.

As most of you know, the bothersome effects of Los Angeles smog have been traced to a photochemical atmospheric reaction between olefinic hydrocarbons and nitrogen dioxide to produce eye irritation, typical vegetation damage, ozone formation, and marked reduction in visibility. Because 85–95% of the olefins in the Los Angeles atmosphere are estimated to come from motor vehicle exhaust, control of exhaust emissions by means of after-burners has been proposed. As a step in this direction, the California State Board of Health has adopted standards for motor vehicle exhaust emissions which call for a reduction of the olefin content of the atmosphere by about 80%.

Now that the authors have called attention to crankcase blowby gases as a significant source of air pollution, several questions arise, which I will try to answer.

1. Will control of crankcase blowby gases markedly reduce the olefin content of the Los Angeles atmosphere? If so, by how much?

In making this compuation, one must take into consideration the extremely poor mechanical condition of the average car in Los Angeles. The average Los Angeles car has four times the hydrocarbon exhaust emission of the average car discussed in the paper, according to the CRC Survey, made in Los Angeles in late 1956. If we assume the olefin content of the exhaust emissions is 45–50% of the total hydrocarbons in both cases, we still have the fourfold difference mentioned.

With regard to blowby, we don't know the relationship between Los Angeles and Detroit cars. As a starting point, let us assume they are the same, inasmuch as the higher exhaust emissions in Los Angeles appear to be a matter of poor maintenance of the electrical system, which should not affect blowby. If we further assume the composition of the blowby is identical to the composition of the fuel, as the paper points out, then the hydrocarbons in Los Angeles blowby would be 23% olefins (the same as the average gasoline, according to the Los Angeles County Air Pollution Control District). In this case, the blowby olefins would be $\frac{1}{4} \times \frac{1}{2}$, or only about $\frac{1}{8}$ of the exhaust olefin emissions ($\frac{1}{4} \times 23/46 = \frac{1}{8}$), or 1/9 of the total. On the other hand, if we assume average blowby emissions in Los Angeles are four times those of Detroit cars, as are the exhaust emissions, then blowby becomes a very important factor, with half as much olefins coming from the crankcase as comes from the exhaust pipe; in other words, one-third of the total.

2. If the California Department of Health had taken blowby emissions into consideration in setting the hydrocarbon emission standard, how would the allowable concentration of hydrocarbons be affected?

Here, we must realize that the Department used two methods to calculate the required reduction in atmospheric olefins: the "rollback" method and the "carbon monoxide indicator contaminant" method. Calculation indicates neither value would have been markedly affected if blowby emissions in Los Angeles were of the same order as those in Detroit.

3. How will these findings affect afterburner development?

As far as the *need* for an afterburner is concerned, there will be little, if any, effect. Even if a different degree of hydrocarbon reduction were required, we would still be faced with the Health Department's carbon monoxide requirement, which is little affected by blowby. A point that might be considered would be piping the blowby continually to an exhaust afterburner. The resulting enrichment could conceivably simplify afterburner design.

4. Are the findings reported of any value outside California?

If photochemical smog exists anywhere in the United States outside of California, and this is doubtful, reduction of blowby emissions would go a long way toward solving the problem. The assumption is made that automobile maintenance elsewhere is similar to that of the five test cars.

In conclusion, it is apparent that more information is necessary on the blowby emissions of representative motor vehicles in California. It is further apparent that blowby is a significant source of atmospheric olefins, whether it is 10 or 50% of the total. Accordingly, the announcement that automobiles sold in California after January 1, 1961, will be furnished with an inexpensive device to control blowby emissions is very encouraging to those of us seeking to alleviate photochemical smog.

Internal Breather Systems One Answer to Problem

— R. J. Templin
Cadillac Motor Car Division, GMC

IN AUGUST, 1959, when the magnitude of the crankcase problem was first clearly established, the authors presented their findings to the Vehicle Combustion Products Committee of the AMA, the socalled VCP. The members of VCP were quick to see both the magnitude of the prob-

lem as well as the strong possibility that something beneficial could be done about it.

A special VCP committee was formed of AMA member companies to pursue the development of a device for eliminating the crankcase emission problem.

The committee's initial development work confirmed that the crankcase gases contained both sufficient oxygen and hydrocarbons to oxidize completely when fed back into the engine induction system. This is in contrast to the problem with exhaust gas, where insufficient oxygen and hydrocarbons are present to initiate or maintain combustion. The fact that the crankcase gases would oxidize meant that internal breathing of the crankcase was a suitable means for controlling this emission.

With the selection of internal breathing as the concept, the choice of one of several types of internal breathing systems was then considered. Two basic types have been used commercially, primarily as a means of providing positive crankcase ventilation. These are:

1. Breathing *above* the carburetor where the air cleaner depression is used to draw the gases out of the crankcase.
2. Breathing *below* the carburetor where intake manifold depression is used as a flow-inducing means.

These systems have both been used since before World War II, but field experience shows that the *above* carburetor system occasionally permits heavy varnish deposits to build up on carburetor parts. This system is still used on some marine engines and a few foreign-built cars, however.

The second method, the *below* carburetor method, is more successful. It has been, and is still being used on large numbers of military engines and larger trucks as a means of keeping dirt and water out of crankcases. It is also used in limited numbers on taxicabs for improving crankcase ventilation. Fig. B shows the "below" carburetor type of system. This is the system chosen for development by the Committee.

The field experience of the truck and taxi manufacturers with this system was surveyed, and this information was exchanged among the member companies. In general, only three field problems of any significance have been found by users of the systems. These are all associated with the flow modulating valve (Fig. C). They are as follows:

1. Gumming or sticking of the valve.
2. Exploding of the crankcase due to induction system backfires forcing flame down the breathing tube.
3. Condensing and freezing of moisture in the system during low temperature operation.

The gumming problem has been obviated by periodic solvent cleaning of the valve; at least once every 10,000 miles is recommended, more frequently in special cases. Placing the valve in a relatively cool location (250 F) is also helpful.

The crankcase explosion problem has been solved by the addition of a backfire stop, either in the modulating valve or in the ventilator line.

The freezing problem is best handled by locating the valve away from cold blasts and water-trapping points.

With this background indicating that systems of satisfactory reliability could be engineered, each company designed and built systems for their respective car engines. As might be expected, in applying the systems to the extremely wide range of operating speeds and loads peculiar to passenger cars, a few *new* problems were unearthed which had not been serious in the heavy vehicle applications. These are as follows:

1. The calibration of the flow control valves is extremely critical to avoid one of three undesirable conditions, namely (a) inadequate flow capacity at some road load speed, (b) excessive flow at idle with resultant poor idle, or (c) inadequate capacity near wide open throttle. Fig. D shows these conditions.
2. The relatively high crankcase airflows into the intake manifold tend to upset mixture control and/or mixture dis-

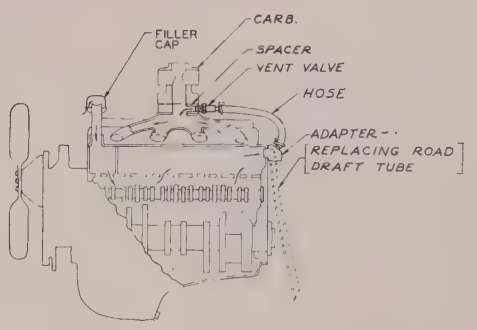


Fig. B — "Below carburetor" breather

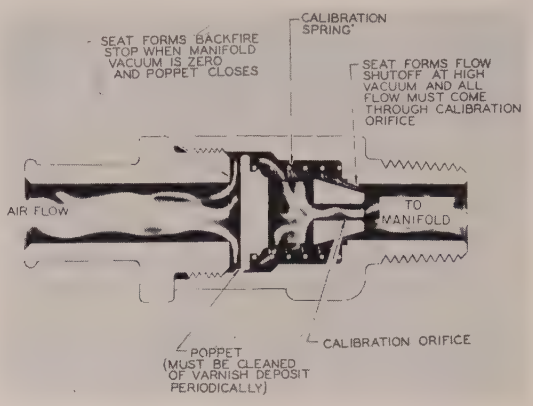

CALIBRATION SPRING
SEAT FORMS BACKFIRE STOP WHEN MANIFOLD VACUUM IS ZERO AND POPPET CLOSES
SEAT FORMS FLOW SHUTOFF AT HIGH VACUUM AND ALL FLOW MUST COME THROUGH CALIBRATION ORIFICE
AIR FLOW
TO MANIFOLD
CALIBRATION ORIFICE
POPPET (MUST BE CLEANED OF VARNISH DEPOSIT PERIODICALLY)

Fig. C — Flow control valve

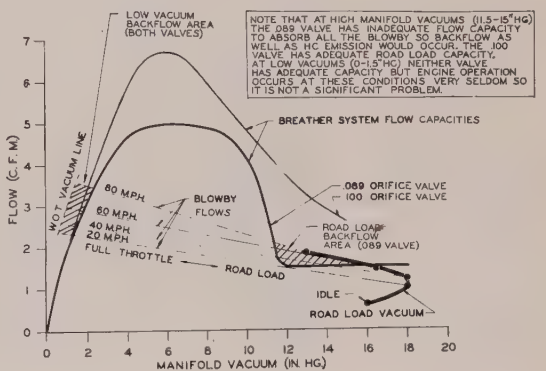

LOW VACUUM BACKFLOW AREA (BOTH VALVES)
NOTE THAT AT HIGH MANIFOLD VACUUMS (11.5–15" HG) THE .089 VALVE HAS INADEQUATE FLOW CAPACITY TO ABSORB ALL THE BLOWBY SO BACKFLOW AS WELL AS HC EMISSION WOULD OCCUR. THE .100 VALVE HAS ADEQUATE ROAD LOAD CAPACITY. AT LOW VACUUMS (0–1.5" HG) NEITHER VALVE HAS ADEQUATE CAPACITY BUT ENGINE OPERATION OCCURS AT THESE CONDITIONS VERY SELDOM SO IT IS NOT A SIGNIFICANT PROBLEM.
BREATHER SYSTEM FLOW CAPACITIES
.089 ORIFICE VALVE
.100 ORIFICE VALVE
ROAD LOAD BACKFLOW AREA (.089 VALVE)
BLOWBY FLOWS
W.O.T. VACUUM LINE
60 M.P.H.
60 M.P.H.
40 M.P.H.
20 M.P.H.
FULL THROTTLE
ROAD LOAD
IDLE
ROAD LOAD VACUUM
FLOW (C.F.M.)
MANIFOLD VACUUM (IN. HG.)

Fig. D — Typical calibration problem in matching flow control valve to engine blowby characteristics

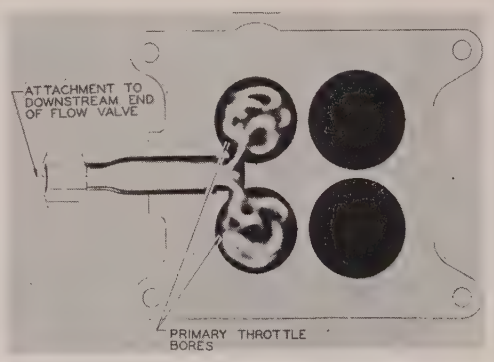


Fig. E — Adapter spacer

tribution at light-load, low-speed conditions. The valve is in effect a manifold leak.

In the case of V-8 engines, this necessitates dividing the air between the manifold branches. A spacer under the carburetor with a T-section passage is one solution to this problem. (Fig. E shows one type.) In some cases modified carburetor calibrations were required to cover lean spots introduced by the airflow from the system. (Fig. F shows the installation on an engine.)

3. Under extreme wide-open throttle conditions, where the available manifold depression is not large enough to pull all the blowby into the system, some of the blowby backs up the oil filler pipe (Fig. B). Loss of oil by entrainment results. This is particularly troublesome on engines where the oil filler pipe (or breather inlet tube) enters very close to the timing chain.

4. The port connecting the breather valve and the intake manifold must be of very low restriction. This is for two reasons. First, the amount of backflow at wide-open throttle is increased markedly by any restriction here, and second, when closing the throttle towards idle, the flow across this restriction creates a pressure drop which prevents the flow valve from sensing full-idle vacuum. The valve cannot come to the full-closed position, therefore, and a poor idle results.

With the possible exceptions of oil entrainment problems, all of the development problems encountered during this work have been solved. In addition, some companies have investigated the effect of these internal breather systems on such problems as: octane requirement increase, oil economy, intake valve and manifold deposit buildup, and crankcase sludge formation.

To date, the systems have not been found to change significantly any of these factors, but it will take extensive field experience to answer completely questions in these areas.

In November, 1959, the special committee's assignment was completed and action was taken by the industry to announce publicly that the systems will be available on all American-built 1961 vehicles sold in California.

Blowby Rate Depends Upon Manifold Pressure

— R. S. Spindt
Gulf Research & Development Co.

DATA OBTAINED in our laboratory supports all of the authors' conclusions, but from another viewpoint.

Their discussion stressed the increasing blowby rate with increasing speed. In our studies, we varied speed and manifold pressure as shown in Fig. G. It is clear that airflow is a direct function of both speed and manifold pressure, while blowby rate depends principally on manifold pressure.

These findings do not disagree with the observation that blowby increases with road load speed, since this speed usually varies directly with manifold pressure. The point is that nobody really drives under constant speed conditions. Traffic driving, even on modern freeways, has resulted in driving habits that causes considerable throttle manipulation. Everytime a driver steps on the throttle, he increases manifold pressure and blowby rate even though he does not hold the condition long enough to increase speed.

Fig. H gives the blowby per cycle in this engine. These

Fig. F — Installation on engine

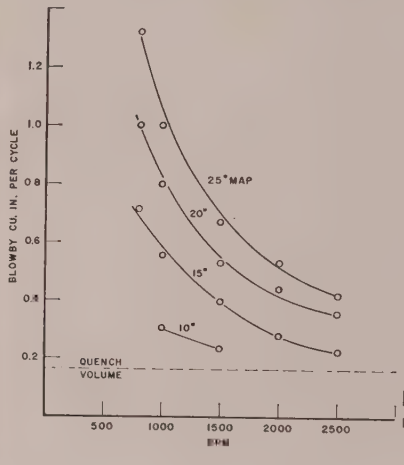


Fig. H — Correlation of blowby rate in volume per stroke with speed and manifold pressure

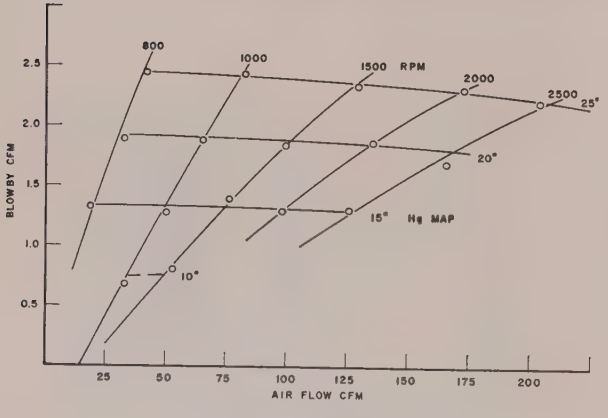


Fig. G — Correlation of blowby with speed and manifold pressure

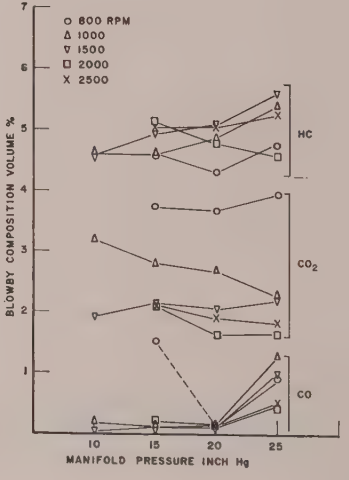


Fig. I — Composition of blowby gases

data, show that while the quench volume, volume A in the paper, can account for most of the blowby gases under some conditions, under others the blowby must also contain gases that were compressed into this area during either compression or combustion. For this reason, the blowby may not be only unburned charge.

The conclusion that the composition of blowby is principally unburned charge has also been substantiated by analysis. Fig. I shows the approximate carbon balance obtained over a range of speed and manifold pressures. Under most operating conditions carbon monoxide is less than 1%, carbon dioxide varies from 2 to 4%, while the carbon on a per atom basis in hydrocarbons is above 4.5–5%. These data are given in terms of readings on Beckman L-B Model 28 analyzers with an ABE cell and a propane calibration for the hydrocarbons. While it is not believed that the use of propane will give a calibration that is absolute for gasoline, it is probable that at these concentrations, the instrument is underestimating the true concentration.

Crankcase Emission Only One Of Many Automotive Pollutants

— Benjamin Linsky
Bay Area Air Pollution Control District

THE AUTHORS introduce the paper by citing three features:

1. Reports recent data on the composition of blowby gases.

2. Evaluates the relative importance of crankcase and exhaust gas hydrocarbon emissions using the best available analytical techniques.

3. Presents and evaluates a method of eliminating the contribution of automotive crankcase gases to air pollution.

I wish to discuss the second of these. (I trust that the data from only five cars will be proved to be truly representative of the general car population, even though past "small population" auto exhaust studies have proved to be misleading.)

The paper contains a pattern of discussing the C_6 hydrocarbons and ignoring such specific hydrocarbons as ethylene. This is unfortunate because ethylene has been proved to be disturbing in two ways:

1. Directly damaging to vegetation.[c]

2. Highly productive of eye irritants and ozone during a few hours of sunlight time, even though some of the other hydrocarbons become "sun-burned" in a shorter time.[d e]

The meaning of these findings seem to me to be as follows:

1. Five cars have been tested and have been found to emit more of some types of hydrocarbons than had been previously recognized.

2. This additional outlet of some types of hydrocarbons is crankcase blowby which can be controlled with relatively simple types of equipment that were developed many years ago.

3. The composition of the additional types of hydrocarbons from crankcase blowby is closely similar to the composition of the gasoline fuel.

4. Future changes in the types of hydrocarbons in the gasoline would change the composition of crankcase blowby if blowby control equipment were not supplied to motor vehicles.

5. The exact ratios of significant blowby hydrocarbons to significant exhaust hydrocarbons is now and undoubtedly will continue to be debated endlessly. The motivation of the discussants guide their selection of criteria and definitions. Those who recognize the need for extensive clean-up refer to each pollutant component separately. Some of those who wish to reduce the degree of clean-up to be legally required refer to groups of pollutants in classes that show the largest apparent benefits from blowby abatement. The arbitrary grouping of specific types of pollutants into arbitrary classes (such as "C_6 hydrocarbons" or "olefins") provides food for debate and confusion rather than enlightenment.

6. It needs to be pointed out that all of the crankcase blowby hydrocarbons might be used to enrich the "feed" to an exhaust afterburner (either direct flame or catalytic in type) during "lean" periods of the driving cycle. This might be less troublesome to carburetors than some of the types of devices described in this paper.[f]

Conclusion: Because of the increased concentration of motor vehicles (and all other human production and use activities) in every metropolitan area in the world, the need is growing for utilizing every practical device and method for reducing each significant air pollutant from motor vehicles. The need is now evident in any metropolitan area that ever experiences poor ventilation accompanied by bright sunshine. This is already a matter of record in California West of the Sierras, in New York City,[g] Washington, D. C.,[h] and in Detroit. The availability of blowby abatement devices is welcome even though it does not reduce the need for effective abatement of the pollutants from the exhaust pipe, the fuel tank, and the carburetor. Information that blowby control will reduce pollutants x and y by 10 or 20% of a motor vehicle's total x and y emissions does not answer the questions about the significance and control of motor vehicle pollutants a, b, and z.

Authors' Closure To Discussion

AMONG THE DISCUSSORS, there appears to be general agreement with our conclusion that crankcase hydrocarbon emissions are a major portion of the automobile's contribution to air pollution. However, the discussors have raised questions in three general areas as follows: (1) The level of blowby flow rate measurements and how they affect our estimate of crankcase emissions, (2) the measured levels of exhaust hydrocarbon concentrations, and (3) the relative amounts of "smog-forming" hydrocarbons in exhaust and crankcase gases.

Concerning blowby measurements, the averaging of the blowby flow rates on our five test cars as carried out by Mr. Stone has no significance. This is because these five cars were not selected to represent any type of car population (Cars C and E were definitely atypical), but rather to illustrate the range in results which has been observed in all of our studies. The conclusion that crankcase and exhaust hydrocarbon emissions are of the same order of magnitude was drawn by considering the blowby flow rates of over 35 cars. The blowby flow rates of these cars are summarized in Fig. J. Also shown for comparison are the average blowby rates at idle and 20 mph presented by Mr. Stone. It may be seen that the overall average blowby rates at idle and 20 mph are very comparable to those he quoted. Accordingly, there appears to be no reason for lowering our estimate that crankcase emissions are approximately 40% of the total engine emission exclusive of carburetor vent losses.

Although the blowby flow rates measured by Dr. Chambers are low relative to the flow rates reported in our paper (many of which are blowby plus ventilation air), his meas-

---

[c] *Smog*, Vol. 60, January, 1960.
[d] "Proceedings of Conference on Chemical Reactions in Urban Atmospheres," by L. H. Rogers. Report No. 15, Air Pollution Foundation, San Marino, Calif., November, 1956.
[e] "Photooxidation of Hydrocarbons in Mixtures Containing Oxides of Nitrogen and Sulfur Dioxide," by E. A. Schuck and G. J. Doyle. Report No. 29, Air Pollution Foundation, San Marino, Calif., October, 1959.
[f] "Smog Burner Fact Sheet," by C. W. Morris. Paper presented at Senate Fact Finding Committee on Public Health & Safety meeting, Los Angeles, Jan. 22, 1960.
[g] *Air Pollution Control News*, Vol. 5, May, 1959.
[h] *Science*, Vol. 129, Jan. 23, 1959, pp. 208–210.

urements fall within the range of results shown in Fig. J. Despite minor differences on specific measurements, it is gratifying to see that Dr. Chambers' data indicate crankcase emissions to be one-third of the total engine hydrocarbon emission compared to our estimated 40%. This agreement is remarkable, considering the number of assumptions necessary to arrive at both of these estimates.

In the second area of questions, exhaust hydrocarbon concentrations are considered. The discussors have referred to the 1956 CRC Survey. Since so many questions have been raised about the effects of instrument calibration and test procedure on the reliability of the CRC Survey results, it is hard to rationalize the use of these data in appraising exhaust gas measurements. In fact, the Survey itself concludes with the often-overlooked statement, "In conclusion, it can be said that the field survey answered a good many questions and raised a good many more; also, that there are still wide variations in reported values and the reasons for these wide variations should be established *before definite conclusions can be drawn with respect to absolute hydrocarbons in automotive exhaust.*"

However, probably of more importance is that hundreds of determinations on vehicles of many makes and models, including many older cars, have indicated that the 1956 CRC survey hydrocarbon concentrations are three to five times too high, except those measured during deceleration. These determinations are not limited to cars run in Detroit by our laboratories and other automobile manufacturers, but include results obtained on Los Angeles cars by the Air Pollution Control District and by the Air Pollution Foundation. It is for these reasons that we consider any arguments based on the 1956 CRC Survey to be of little value.

The third area of questions is concerned with the amount of "smog-forming" compounds in blowby and exhaust gases. In our paper we have considered total hydrocarbons from both sources without regard for types of hydrocarbons. In his discussion, Dr. Faith states that only olefinic hydrocarbons lead to photochemical smog formation. Moreover, Mr. Maga indicates this assumption about olefins was used in developing the California standards for exhaust emission from motor vehicles.

To answer the question on the relative olefin content of blowby and exhaust gases, we recently obtained some composition data by gas chromatography. The blowby and exhaust samples were taken at 30 mph-road load while operating one car using three fuels of varying olefin content.

Chromatographic analyses were made essentially as described in the paper with one important addition to the technique. This addition was the use of a mercuric perchlorate—perchloric acid column which permits a complete separation of paraffin and olefin hdrocarbons.

The result of analyses of both blowby and exhaust gas samples (olefin content expressed as a volume percent of the total hydrocarbons present) are lisited in Table G. Also shown are the usual FIA determinations of olefin content of the fuels. The results show that the olefin content of the hydrocarbon portion of blowby is very similar to the olefin content of the fuel. The olefin content of the hydrocarbon portion of exhaust is much greater than the olefin content of the fuel.

However, there is little justification for considering all olefins as a class, and stating that total olefin content is an indication of propensity for smog formation. For example, Air Pollution Foundation Report 29 presents the "smog-forming" reaction rates of many of the more common hydrocarbons. The data in this report[e] show that the first member of the olefin family, ethylene, has an extremely low reaction rate. In fact, its rate is as low as the reaction rates of isooctane and toluene. Therefore, if this reaction rate is an indication of the smog-forming potential of a compound as suggested by the report, this data supports the *exclusion* of ethylene from the undesirable "smog-forming" olefins. If the concentration of ethylene is subtracted from the total olefin content of the exhaust and blowby gases, the resulting concentrations should be more indicative of the smog-forming tendency of the two hydrocarbon sources than the total olefin concentration, per se.

In Table H the olefin content excluding ethylene expressed as a volume percent of the total hydrocarbons present are listed for the blowby and exhaust samples previously shown in Table G. The resulting olefin percentages for exhaust gases are reduced markedly because a large portion of the exhaust olefins is ethylene. Consequently, the hydrocarbons present in blowby gases have a larger percentage of "smog-forming" olefins than the hydrocarbons in exhaust gases. This suggests that a consideration of specific compounds in these gases adds to, rather than subtracts from, the importance of crankcase hydrocarbon emissions in the problem of smog formation.

In summary, it is evident that more information is needed in several areas to augment the findings of our paper. Certainly many more cars should be studied in order to establish firmly the relative magnitude of crankcase and exhaust hydrocarbon emissions for a typical car population. Further studies should be made on the specific hydrocarbons present in blowby and exhaust gases. In addition, information concerning the compounds and reactions leading to smog formation will be necessary to completely assess the importance of crankcase gases in the problem of automotive air pollution.

Table G — Olefin Content

	Volume[a] % Fuel	Volume % of Total[b] Hydrocarbons Present	
		Blowby	Exhaust
Fuel A	11	10.1	39.8
Fuel B	8	11.9	35.7
Fuel C	38	36.9	52.9

[a] FIA analyses.
[b] Gas chromatographic analyses.

Table H — Olefin Content Excluding Ethylene

	Volume[a] % Fuel	Volume % of Total[b] Hydrocarbons Present	
		Blowby	Exhaust
Fuel A	11	9.5	6.8
Fuel B	8	11.6	8.6
Fuel C	38	36.2	15.6

[a] FIA analyses.
[b] Gas chromatographic analyses.

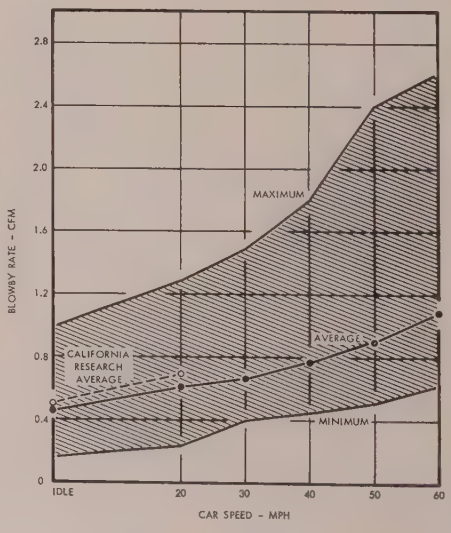


Fig. J — Summary of blowby rates measurements on 38 cars

Material and Structural Damping for Vibration Control

B. J. Lazan
University of Minnesota

THE INCREASING importance of near-resonant vibrations and the role of system damping in their control is discussed. The various component parts of system damping are classified and analyzed within the framework of:

1. Hysteretic damping within the structural materials.
2. Structural damping associated with (a) interface slip or Coulomb friction, and (b) shear strain in an adhesive layer at an interface.

Each of these mechanisms is analyzed to emphasize the factors important in the utilization of damping as an engineering property. Particular attention is paid to analytical concepts for maximizing the shear damping in an interface adhesive by design optimization procedures.*

THE CURRENT trend towards lighter weight, reduced factors of safety, and higher speeds has greatly intensified the dynamic loading encountered by a structure in service. This same trend has also brought the natural frequencies of structural components and assemblies into a region where near-resonant vibrations are becoming increasingly prominent in service. Therefore, the deleterious effects of the fatigue stress associated with near-resonant vibration has become a major cause for service failure.

In former years resonance conditions could often be avoided by properly controlling the natural frequencies in a system — that is, by adjusting the stiffness and mass factors in the structure such as to separate the natural and the exciting frequencies safely. However, this approach is no longer effective in many situations, particularly in aero-space vehicles, because an excessive weight penalty is often involved. Furthermore, random excitation, either of mechanical or acoustical origin, is becoming more prevalent in service. For example, the typical noise spectrum shown in Fig. 1 for a jet engine[1] contains a range of frequencies which is wide enough to encompass most of the resonance frequencies encountered in an aircraft structure. Thus, whatever the natural frequency of the structural components and assemblies may be within the wide range of frequencies shown in Fig. 1, resonant vibrations can still be excited.

Experiences in recent years have made it abundantly clear that resonant vibrations can no longer be avoided by clever design. Modern structures, particularly aero-space vehicles, must be so designed as to withstand the resonant excitations characteristic of service. This paper is concerned with the factors involved in (1) the analysis of resonance amplification, (2) the fatigue stress associated with this condition (resonance fatigue), and (3) the importance of damping in the design of structures for high resonant fatigue strength. Since the factors important in conventional fatigue

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 7, 1959.

1 "Near Field Jet Noise," by M. O. W. Wolfe. NATO-AGARO Report 112, April, 1957. (Available through NASA.)

strength (stress concentration, loading range, and the like) are amply covered in other papers, this paper shall be concerned with only those factors which are of unique interest in resonance fatigue.

By way of introduction and in order to clarify the role of damping in resonance fatigue the elementary resonance curves for a simple single-degree-of-freedom system excited by a simple *sinusoidal* force shall be reviewed.

Referring to Fig. 2, mass M is attached to a light structural member A which is held at its top by the support and cover plates B. The *exciting force* imposed on the mass is $F_g \sin \omega t$ and the force felt by the support, the *induced force,* is $F_d \sin \omega t$. At very low exciting frequencies ω, well below the resonant condition, induced force F_d is equal to or only slightly larger than exciting force F_g. However, as the exciting frequency approaches the natural frequency of the system (ratio of two frequencies shown in abscissa approaches unity) a vibration amplification is observed, the magnitude of which (A_v in Fig. 2) depends on both the proximity to resonance and the damping in the system. The maximum amplification occurs at resonance when the vibration amplification factor A_v is called the resonance amplication factor A_r. Thus, under sinusoidal excitation:

$$A_v]_{\phi=90°} = A_r = F_d/F_g \qquad (1)$$

where:

ϕ = Vibration phase angle = 90 deg at resonance

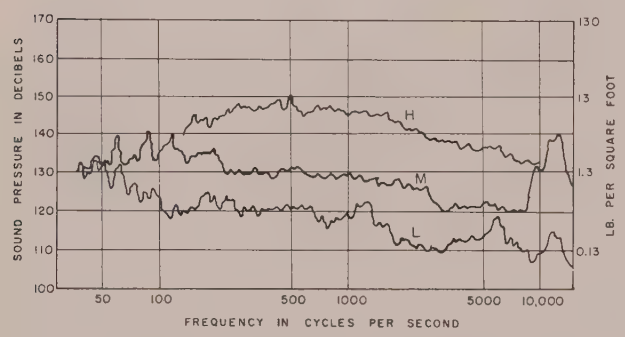


Fig. 1 — Frequency content of typical jet exhaust noise at high (H), medium (M), and low (L) exhaust velocities

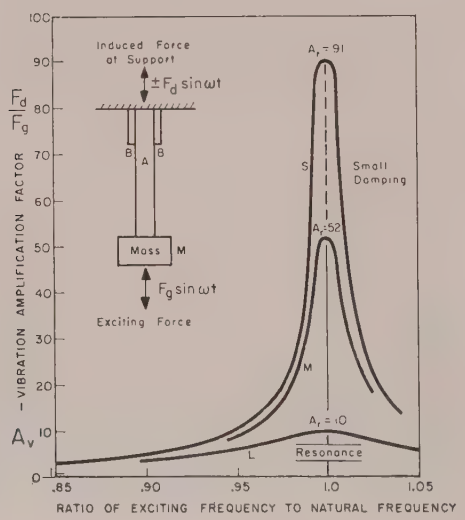


Fig. 2 — Effect of system damping on vibration amplification

For the case illustrated, the resonance amplification factor A_r is 91 for a system with small damping (S curve), 52 for medium damping (M curve), and 10 for large damping (L curve). The great importance of the system damping in defining resonance amplification and controlling the associated resonance fatigue problems is thus apparent.

It can be shown[2] that for a linear system under sinusoidal excitation:

$$A_r = 2\pi \frac{W_e}{D_o} \qquad (2)$$

where:

W_e = Total elastic or strain energy in the system partaking in the vibrations at its maximum deflection, in.-lb

D_o = Total damping energy dissipated by the system per cycle of vibration, in.-lb per cycle

Combining Eqs. 1 and 2:

$$F_d = 2\pi \frac{W_e}{D_o} F_g \qquad (3)$$

The above equation may be interpreted as follows in a fatigue analysis problem. Induced force F_d is the force actually felt by the structure (member A in this case) and provides a basis for the conventional fatigue analyses. That is, knowing the induced force and such structural details as shape of members, notch geometry, and the like, one can compute the fatigue stress at the maximum locations and determine fatigue life. In the nonresonance problem, the induced force is equal to the exciting force F_g so the conventional fatigue analyses may be applied directly to the exciting force. However, in the resonance fatigue problem the exciting force may be greatly amplified, since W_e/D_o in Eq. 3 may be a large number. Thus, in a resonance fatigue analysis of a system under sinusoidal excitation one is concerned not only with the conventional fatigue problems but also with the degree of amplification. As shown by Eq. 3, the total damping D_o in a system is an important parameter in such an analysis.

The total damping is equally important in a system under random excitation of acoustical origin. Even though such excitation is (1) produced by fluid pressure rather than structural forces, (2) not periodic, and (3) variable in amplitude at any discrete frequency (the resonant frequency) rather than of constant amplitude, the role of damping in controlling random resonant vibration is similar to that indicated above for periodic resonant fatigue.

In spite of its great importance in a variety of types of resonant fatigue problems damping has received relatively little attention to date. Several years ago a program of study on damping was initiated at the University of Minnesota, first on materials damping and more recently on structural interface effects. The purpose of this paper is to review some of the results of this study.

Damping and Damping Mechanisms

The term damping as used in this paper defines the *energy dissipation* properties of a material or system under cyclic stress. In most cases a direct

[2] "Effect of Material Damping and Stress Distribution on Resonant Fatigue Strength of Parts," by E. Podnieks and B. J. Lazan. WADC TR 55-284, August, 1955.

conversion of mechanical energy to heat is involved. This definition specifically excludes such *energy transfer* devices as dynamic absorbers or so-called dynamic dampers." Within the context of this definition energy must be absorbed and dissipated within the specified system before the term damping is applicable.

Damping may be classified in various ways. For convenience in this paper damping is broken down into two major headings which shall be identified in this discussion as: material damping and system damping.

Material damping, sometimes called internal friction, internal damping, or hysteretic damping, is related to the energy dissipation in a *volume of macrocontinuous media*. The term macrocontinuous is intended to exclude the damping in a configuration originating at interfaces between recognizable parts, yet include the types of micro- and submicrointerface effects which might constitute an important mechanism in the volume or bulk damping of materials not homogeneous on a microscopic or submicroscopic scale. In general, material damping is associated with the energy dissipation which takes place when a more or less homogeneous volume is subjected to cyclic stress and the damping mechanisms are associated with the internal microstructure of the material. The two types of materials of interest in structural analyses to be discussed in this paper are (1) structural metals and nonmetals and (2) viscoelastic adhesives.

Whereas material damping occurs in a volume of a macrocontinuous medium *system damping* involves configurations of distinguishable parts or the interaction among various phenomena. Among the types of systems in which damping under cyclic stress may be important are:

1. Structural systems in which energy is dissipated in various types of joints, interfaces, or fasteners.
2. Electromechanical systems in which energy conversion and dissipation may take place through the interaction between electrical or electromagnetic phenomena and physical bodies. Examples of this type of energy loss is the system damping associated with magnetic hysteresis and eddy currents.
3. Hydromechanical and acoustical systems in which damping occurs through fluid flow. Acoustical damping and radiation, oil flow through orifices, and dashpot effects are examples of this type of damping.

Since this paper is concerned primarily with structural damping further reference will not be made to electromechanical, hydromechanical, or acoustical damping.

In order to clarify the types of damping of interest in aero-space vehicles it is desirable to ask the following question: Beyond the damping energy dissipated internally by the materials in the structure what is the principal mechanism by which structural joints dissipate energy?

The unique characteristic of a joint insofar as damping is concerned is its interfaces or mating surfaces which are maintained on contact. Thus, it is the damping associated with various interface effects that should receive close scrutiny.

For purposes of this paper, two types of interfaces shall be identified: (1) a dry or lubricated contact surface (metal-to-metal[3] contact or metal-to-liner-to-metal contact), or (2) an adhesive-type interface (metal-to-adhesive-to-metal joint). Next, let us consider the simple types of relative motions which can take place between mating surfaces at an interface to produce damping effects. A review of the general behavior of typical joints in typical load environments indicates that two types of relative motion should be considered: a separation of mating surfaces (motions perpendicular to the interface), and interface shear effects (relative motions in the plane of the interface). Of these two types of motion the one which appears to offer the greater potential for dissipating energy is relative shear at an interface. For example, if member *A* in the system shown in Fig. 2 is vibrated axially, the cyclic axial strain in member *A* will tend to produce a cyclic shear displacement (slip) at the interfaces between member *A* and cover plates *B*. This slip provides a mechanism for dissipating energy.

For the case of dry interfaces (metal-to-metal contact) coulomb friction provides the mechanism for dissipating energy under cyclic shear displacement. Several types of joints have been investigated considering this mechanism. It is found as discussed later, that this mechanism is capable of dissipating very large energy if the joint design is properly optimized.[4, 5, 6]

It is generally found, however, that a design optimized for maximum dry slip damping may sometimes develop serious fretting and corrosion effects in the interface regions subjected to large cyclic slip. Such interface surface deterioration may not only cause the joint to drift from optimum conditions but may also initiate fatigue cracks, the very thing that high damping in a system exposed to resonant vibration is intended to mitigate. In short, the cure may lead to a condition worse than the original disease.

In view of this difficulty with dry interface slip consideration was given to the possibility of lubricated interfaces[7] and inserts of plastic and other types of "nonfretting" materials at interfaces.[8] However, it was still difficult to avoid fretting and maintain optimum conditions for maximum damping in practical joints which dissipate large damping. This led to the consideration of an adhesive-type interface of sufficient thickness to permit the relative shear motions between metal surfaces to be absorbed as shear strain within the adhesive itself (no relative motion between adhesive and the metal adherents). This is the principle of sound deadening tape, which has been in use for about a decade. However, there was until recently very little interest in the utilization of adhesive damping in structural joints. The potential contribution of this approach became more apparent recently when it was

[3] The term "metal" is used here for conciseness. Actually the term structural material, which includes not only metals but plastics, ceramics, and the like is implied.

[4] "Effect of Material and Slip Damping on Resonance Behavior," by B. J. Lazan and L. E. Goodman. Paper presented at ASME Symposium on Shock and Vibration, Applied Mechanics Division, June, 1956.

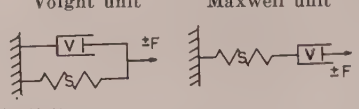
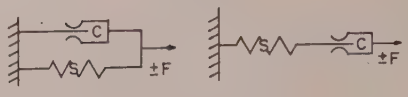
[5] "Material and Interface Damping," by B. J. Lazan and L. E. Goodman, Ch. 36 of "Handbook of Shock and Vibration Control," by C. Harris and C. Crede. Pub. by McGraw-Hill Book Co., New York.

[6] "Study of Structural Damping of Simple Built-Up Beam with Riveted Joints in Bending," by T. H. H. Pain. Technical Report, Contract N5ori-07833 to ONR Aeroelasticity and Structures Laboratories, M.I.T., May, 1954.

[7] "Frictional Damping and Resonant Vibration Characteristics of Axial Slip Joint," by J. H. Klumpp and B. J. Lazan. WADC TR 54-64, March, 1954.

[8] "Effect of Damping Compound on Jet-Efflux Excited Vibration," by D. J. Mead. Reports ARC 20,155 and 20,263 (also ARC 19,870), Aeronautical Research Council, England, 1958.

Table 1 — Classification of Types of Hysteretic Damping of Materials

Name Used Here	Types of Material Damping			
	Dynamic Hysteresis	Static Hysteresis		
Other Names	Viscoelastic, rheological, and rate-dependent hysteresis	Plastic, plastic flow, plastic strain, and rate-independent hysteresis		
Nature of Stress-Strain Laws	Essentially linear. Differential equation involving stress, strain, and their time derivatives	Essentially nonlinear, but excludes time derivatives of stress or strain		
Special Cases and Description	*Anelasticity.* Special because no permanent set after sufficient time. Called "internal friction"			
Simplest Representative Mechanical Model	Voight unit Maxwell unit  anelasticity			
Frequency Dependence	Critically at relaxation peaks	No, unless other mechanisms present		
Primary Mechanisms	Solute atoms, grain boundaries. Micro- and macrothermal and eddy currents. Molecular curling and uncurling in polymers		Magnetoelasticity	Plastic strain
Value of n in $D = JS^n$	2		3 — up to coercive force	2-3 up to σ_L 2 to >30 above σ_L
Variation of η with Stress	No change, since $n-2=0$		Proportional to σ since $n-2=1$	Small incr. up to σ_L Large incr. above σ_L
Typical Values for η	Anelasticity: <0.001 to 0.01 Viscoelasticity: <0.1 to >1.5		0.01 to 0.08	0.001 to 0.05 up to σ_L
Stress Range of Engineering Importance	Anelasticity — low stress Viscoelasticity — all stresses		Low and medium. Sometimes high	0.001 to >0.1 above σ_L Medium and high stress
Effect of Fatigue Cycles	No effect		No effect	No effect up to σ_L Large changes above σ_L
Effect of Temperature	Critical effects near relaxation peaks		Damping disappears at Curie temperature	Mixed. Depends on type of comparison
Effect of Static Preload			Large reduction for small coercive force	Either little effect or increase

found that the damping capacity of viscoelastic adhesives (as a material) in shear in very large as discussed later. Furthermore, it was also found that viscoelastic adhesives can withstand very large cyclic shear strain without deterioration. This combination indicates that a properly optimized adhesive joint (thickness, for example, adjusted properly in accordance with its stiffness and other properties) can dissipate very large damping energy. Furthermore, this damping mechanism not only avoids corrosion and fretting problems but also simplifies the problem of maintaining optimum conditions during service.

Material Damping

Materials are not perfectly elastic even at very low stress levels. Inelasticity in materials manifests istelf in a variety of different ways. Under cyclic stress, for example, inelastic behavior takes the form of a stress-strain hysteretic loop, such as that illustrated in Fig. 3. Although such loops are always present at stress of engineering interest they are often too narrow to be observed by conventional methods. The shape of the loop depends on the damping mechanism operative.

Many different types of inelastic mechanisms and hysteretic phenomena have been identified. For purposes of this paper the damping classification given in Table 1 is appropriate. The various damping phenomena and mechanisms may be identified under two main headings: dynamic hysteresis and static hysteresis.

Dynamic hysteresis is sometimes identified as viscoelastic, rheological, and rate-dependent hysteresis. It is observed in materials having essentially linear stress-strain laws which are describable by a differential equation containing stress, strain, and time derivatives of stress or strain. These differential equations need not be linear. Furthermore, they can include terms which allow for permanent set such as *OB* after cycle *OAB* in Fig. 3.

One important type of dynamic hysteresis, a special case which Zener[9] has labeled anelasticity, does not include provisions for permanent set after a long time. This means that if the load is suddenly removed at point *B* in Fig. 3, after cycle *OAB,* strain *OB* will gradually reduce to zero as the specimen recovers (or creeps negatively) from point *B* to point *O*.

The terms anelasticity and internal friction (damping in anelastic materials) has been reasonably well accepted by physical metallurgist for over a decade. However, the more general types damping identified here as viscoelastic and rheological hysteresis do not have a well-accepted name.

A distinguishing characteristic of internal friction and the more general case of viscoelastic damping is its dependence on time derivative effects. Thus, the hysteresis loops tend to be elliptical in shape rather than pointed as in Fig. 3. Furthermore, the loop area is definitely related to the dynamic or cyclic nature of the loading and the area of the loop is dependent on frequency. In fact, the stress-stress curve, for an ideally viscoelasticity material becomes a single value curve (no hysteretic loop) if the cyclic stress is applied slowly enough to allow the material to be in complete equilibrium at all times (oscillation period very much longer than relaxation times). No hysteretic damping is produced by these mechanisms if the material is

[9] "Elasticity and Anelasticity," by C. Zerner. Pub. by University of Chicago Press, Chicago, 1948.

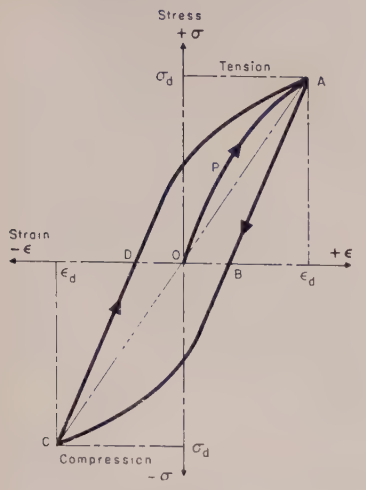


Fig. 3 — Typical stress-strain hysteresis loop for material under cyclic stress

subjected to essentially static loading. Stated differently, the static hysteresis is zero.

Static hysteresis, in contrast with dynamic hysteresis, involves stress-strain laws which are insensitive to time, strain or stress rate, or other derivatives. In a material in which static hysteresis dominates, the value of strain is attained almost instantly for each value of stress and prior stress history (direction of loading, amplitudes, and the like), independent of loading rate. Under cyclic loading pointed loops similar to Fig. 3 are formed and if the stress is reduced to zero (point B) after cycle OAB, then OB remains as a permanent set or residual deformation. Furthermore, the shape of the hysteretic loop is independent of frequency.

The two principal mechanisms which lead to static hysteresis are magnetostriction and plastic strain. Therefore, Table 1 shows two headings under static hysteresis, not only to identify the mechanisms operative, but also to provide a name for describing their characteristic behavior discussed later.

Table 1 also shows the simplest representative mechanical models for each of the behaviors classified. In these models S is a spring having linear elasticity (linear and single-valued stress-strain curve), V is a linear viscous dashpot which produces a resisting force proportional to velocity, and C is a coulomb friction unit which produces a constant force whenever slip occurs within the unit, the direction of the force being opposite to the direction of motion. More sophisticated models have been found to predict reliably the behavior of some materials,[10] particularly polymeric materials.

The various inelastic mechanisms indicated in Table 1 have been discussed in prior publications[9, 10, 11] and will not be reviewed in this paper. However, in order to facilitate comparisons to be made later, with other damping mechanisms, it is desirable to indicate magnitudes of damping which can be expected from structural materials. Such data are given in the bottom half of Table 1, which is self-explanatory. The units and symbols used are defined below:

σ_d = Induced stress in material, psi
D_d = Specific damping energy of a material at induced stress σ_d, in.-lb per cu in. per cycle
E = Modulus of elasticity, psi

$$\eta = \text{Loss factor of material} = \frac{E}{\pi}\frac{D}{\sigma_d^{\,2}} \qquad (4)$$

σ_e = Fatigue strength of the material, psi
σ_L = Limiting stress approximately 80% of σ_e, psi

Various criteria for comparing the damping properties of material have been used.[2, 11] It has been shown that a convenient and significant method for comparing the damping properties of structural materials is to plot damping energy D_d versus ratio of induced stress to fatigue strength (ratio σ_d/σ_e). Such a pilot is shown in Fig. 4 for a variety of common structural materials. For the large group of structural materials which were not particularly selected for high damping (excludes materials having large magnetoelastic or plastic strain damping) and for a variety of test conditions[11] the data lie within a fairly well-established hand shown in Fig. 4. The approximate "geometric" mean curve is shown as the dot-dash in Fig. 4. Even though this is a two-segment line it can be defined with sufficient accuracy by the single equation:

$$D = (\sigma_d/\sigma_e)^{2.3} + 6(\sigma_d/\sigma_e)^{8} \qquad (5)$$

Also shown in Fig. 4 for comparison purposes are four materials having especially high damping. Materials 1 and 2 are the magnetoelastic alloys Nivco 10 and 403 alloy. Nivco 10[12] retains its high damping up to the stresses shown (data not available at higher stresses). However, the 403 alloy reaches its magnetoelastic peak at a stress ratio of approximately 0.2 and increases less rapidly beyond this point, up until plastic strain damping becomes dominant (at stress ratio of approximately 0.8), beyond which damping increases very rapidly. By contrast, Material 3, a manganese copper alloy[13] with large plastic-strain damping, retains its high damping up to and beyond its fatigue strength. Material 4 is a "typical" viscoelastic adhesive (Tape No. 466, Minnesota Mining & Mfg. Co.) for which a cyclic shear strain of unity is assumed to lie within the fatigue strength. (Experiments show that a shear strain well above unity does not cause deterioration in the adhesive even after millions of cycles.)

Since these damping data are plotted to a logarithmic scale, the superiority of the high damping materials is not dramatically revealed by Fig. 4. However, Nivco 10, for example, has a damping 30 times as large as the average structural material in the stress range shown in Fig. 4, and the viscoelastic damping is over 10 times as large as Nivco. These observations on viscoelastic adhesives will be utilized later.

System Damping and Materials Damping

Interface Slip — It has been shown both theoretically and experimentally[4, 5, 6] that interface slip has

[10] "Mechanical Behavior of High Polymers," by T. Alfrey, Jr., pp. 191–192. Pub. by Interscience Pub. Co., New York, 1948.
[11] "Fatigue Failure under Resonant Vibration Conditions," by B. J. Lazan. Pp. 36–76 of "Fatigue," pub. by ASM, 1954.
[12] "Damping of Flexural Vibrations by Alternate Visco-elastic and Elastic Layers," by E. E. Ungar, D. Ross, and E. M. Kerwin, Jr. Pub. by Bolt Beranek and Newman, Cambridge, Mass., Report No. 640, July, 1959 (to be published as WADC Report). See also "Damping of Flexural Waves by Constrained Viscoelastic Layer," by E. M. Kerwin, Jr. *Journal* of Acoustical Society of America, Vol. 31, July, 1959, pp. 952–962.
[13] "Damping Properties of Cast Magnesium and Manganese Copper Alloy Proposed as High Damping Material," by P. Torvik. Appendix 72fg, Status Report 58-4, by B. J. Lazan, University of Minnesota, Dec. 13, 1958.

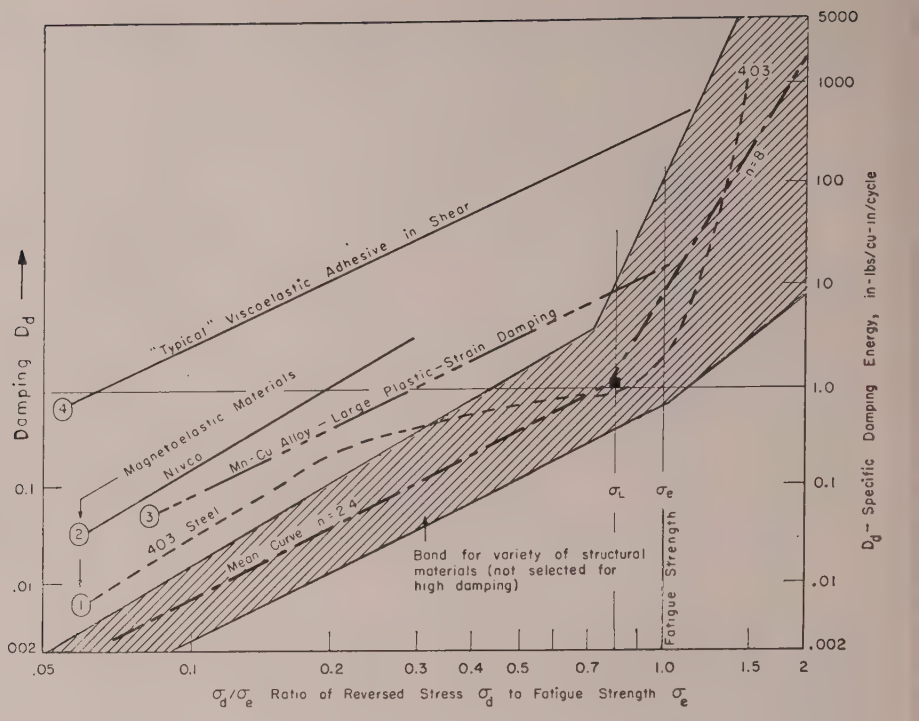


Fig. 4 — Range of damping properties for variety of structural materials

the potential for dissipating large damping energy. Still, engineering designers have made little or no use of slip damping. This has been due partially, and in many cases justifiably, to the fear of fretting and fretting fatigue which may in some cases be a greater evil than low damping. However, in many cases the failure to realize the potential in slip damping is due to unfamiliarity with its basic nature and design parameters. Although some slip damping is present in practically all structures this damping mechanism has not been adequately studied in the past.

Reasonably general equations for slip damping have been developed recently.[5] However, since the purpose of this paper is to convey general damping concepts for engineering guidance only one special case shall be considered to illustrate the important parameters.

The case to be discussed is the bileaf cantilever beam *AB* under uniform pressure *P* as shown in Fig. 5. At very small values of exciting force F_g (or at very large interface pressure) the friction between beams *A* and *B* is sufficient to prohibit slip, and the bileaf behaves as though *A* and *B* were one solid beam without interface. However, if the exciting force is increased (or if pressure *P* is reduced) slip will occur causing an offset at the end of the beams labeled *S*. Under cyclic exciting force, slip *S* at the end of the beam and also along the entire interface will also be cyclic and provide a mechanism for dissipating damping energy. This energy is a function of both the slip and shear stress at the interface. Under large pressure *P* the shear stress is large but slip small; whereas under small pressure *P* the shear stress small but the slip is large. Since damping energy depends on a product of a shear stress function and a slip function, maximum damping will occur at some intermediate or optimum pressure. The equation for defining the damping energy dissipation associated with this interface slip have been developed and are plotted in Fig. 5B for one special joint considered theoretically and experimentally.[4] This figure shows the total damping D_0[8] due to interface slip only as a function of interface pressure *P* and ratio of induced stress σ_d in the beam to its fatigue strength σ_e. The maximization of damping at an optimum pressure of 80 psi is clearly shown in this figure. Another feature of this figure (to be contrasted later with material damping) is that the damping does not change abruptly with stress.

The relative magnitudes of slip and material damping cannot generally be compared directly, due to the basic differences in the mechanisms involved. The material damping in a part depends not only on its material composition and constitution but also on stress amplitude, the volume-stress function of the part, and other factors discussed in prior publications.[11] Slip damping at an interface depends on another set of parameters, which include the coefficient of friction, pressure, shear stress, and strain distribution. Furthermore, material damping occurs throughout the volume of a part (a volume integral), whereas slip damping occurs at an interface surface only (a surface integral). Thus, the two types of damping are not directly comparable. Parts and joints can easily be conceived in which either slip or material damping will dominate. Nevertheless, in order to clarify the important parameters in damping it is desirable to make simplified comparisons between material and slip damping. In these comparisons only two important parameters shall be considered — interface pressure and maximum stress. Other factors such as member geometry, coefficient of friction, and types of vibration, are held constant.

For comparison purposes material damping for the same beam assembly, computed for a "representative" material defined by Eq. 5 is shown in Fig. 5A. (See also Fig. 4.) The main features this figure

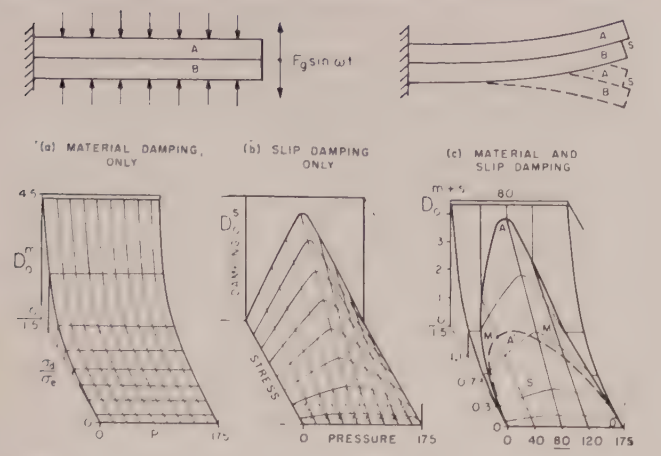


Fig. 5 — Material and slip damping of bileaf cantilever beam under uniform interface pressure

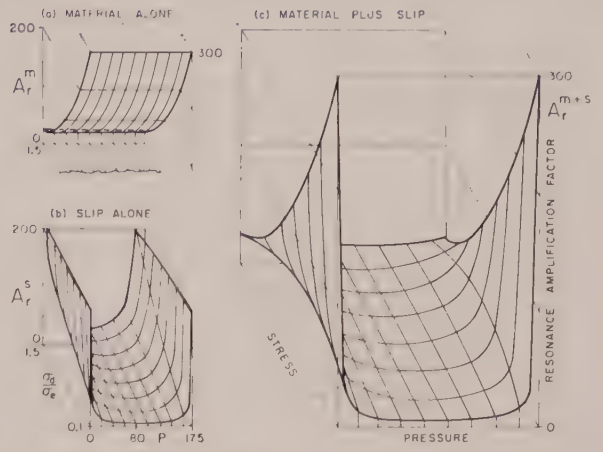


Fig. 6 — Resonance amplification factor of bileaf cantilever beam considering both material and slip damping

has are: damping increases rather abruptly at a stress in the vicinity of the fatigue limit, and interface pressure does not significantly affect material damping.

Both material and slip damping are plotted in Fig. 5C, the line OAD indicating the intersection between the two surfaces. The projection of this line on the basal plane *OA'D* indicates the combinations of stress and pressure in which each type of damping dominates; in region *S* at near optimum pressure and at intermediate stress the slip damping is the larger, whereas in region *M* at extremely high or low pressure or at high stress the material damping dominates.

A second criterion for judging resonance behavior is resonance amplification factor A_r, which varies inversely with damping as shown in Eq. 2. Considering material damping only A_r^m decreases rapidly with increasing stress as shown in Fig. 6A, from 300 at 10% of the fatigue strength to 12 at 150% of the fatigue strength. Interface pressure is not a significant factor in material damping, as observed previously. Considering next slip damping alone both interface pressure and induced stress are important factors as shown in Fig 6B. A minimum A_r^s is attained under the optimum pressure of 80 psi and at low induced stress. If both material and slip damping are included in the computation, then the resultant resonant amplification factor curve A_r^{m+s} combines the features of both curves (Fig. 6C).

Finally, a criterion based on the permissible exciting force F_g (Eq. 3) which produces various amplitudes of cyclic stress at resonance is shown in Fig. 7. In order to place the F_g criterion on a dimensionless basis, the ratio F_g/F_e is plotted, F_e being that alternating force which if applied statistically will produce a stress σ_e (the fatigue strength) in the beam. Thus, F_e is an indication of the conventional fatigue strength of the beam and F_g at a value of $\sigma_d/\sigma_e = 1$ indicates the resonant fatigue strength. In a system with low damping the exciting force F_g can only be a small fraction of force F_e (ratio F_g/F_e small) before dangerously high fatigue stress is developed. The shape of the exciting force curves shown in Fig. 7 are essentially the same as the damping curves shown in Fig. 5. Force F_g is large at low stress under optimum pressure (where slip damping

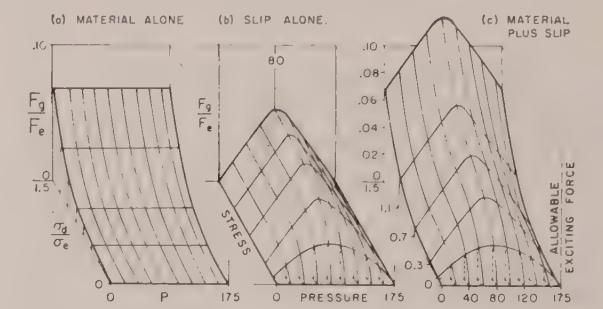


Fig. 7 — Allowable exciting force at resonance for bileaf cantilever beam considering both material and slip damping

is large) and at high stress (where material damping is large).

Shear in Interface Adhesive. As discussed previously the fretting and surface corrosion associated with dry slipping (and developed even in lubricated surfaces) under optimum pressure has led to the study of another damping mechanism: shear hysteretic damping in a viscoelastic adhesive layer located at the interface between rigid members. No shear slip between adhesive and adherents is assumed in such a configuration and all of the relative shear displacement between adherents is taken up as shear strain within the adhesive.

The shear properties of viscoelastic materials are generally specified in the complex notation:

$$G_r = G_1 + iG_2 \qquad (6)$$

where:

G_r = Total, resultant, or absolute modulus of rigidity, or complex rigidity
$= [G_1^2 + G_2^2]^{1/2}$
G_1 = Real part of modulus, or *storage modulus*
G_2 = Imaginary part of modulus, or *loss modulus*

The specific damping energy D, in terms of the above notation is:

$$D = \pi G_2 \gamma^2 \text{ in.-lb/cu in./cycle} \qquad (7)$$

where:

γ = Unit shear strain amplitude in the adhesive layer

For illustrative purposes let us consider one rather

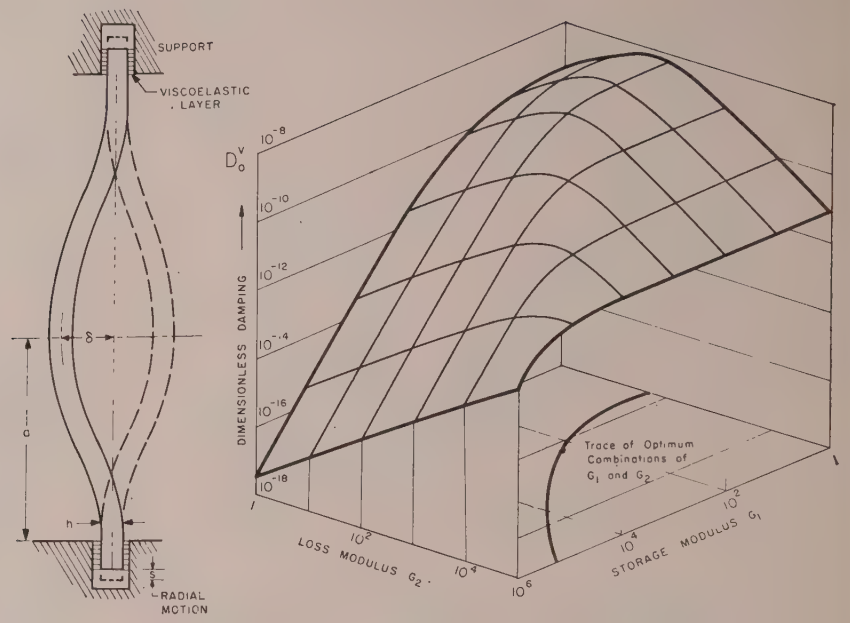


Fig. 8 — Energy dissipation at supports of circular plate for different properties of adhesive supports

typical adhesive[14] and compute its energy dissipation potential as compared to structural metals. At room temperature at 50 cps sinusoidal frequency this adhesive has the following moduli:[15] $G_1 = 95$ psi, $G_2 = 110$ psi, and $G_r = 145$ (moduli values are about twice as high at 250 cps). Furthermore, this adhesive can withstand large cyclic shear safely; shear strains greater than unity at low frequency cause no apparent deterioration in millions of cycles. At unit shear strain damping $D = \pi(110)\ (1)^2 = 345$ in.-lb/cu in./cycle. This damping energy is about 100 times larger than the D values for a "typical" structural metal at its fatigue limit. Furthermore, through proper design, as illustrated later, the adhesive layer can be subjected to uniform shear strain, and thus be capable of dissipating maximum energy at all locations. By contrast, most structural parts have rather severe stress gradients and only a very small percentage of their total volume is at maximum stress and dissipating maximum unit damping energy. (See Lazan[11] for quantitative treatment of this factor through use of volume-stress functions.) It is apparent that viscoelastic adhesives, if properly used in a configuration can dissipate very large damping energy.

In order to clarify the general concepts involved in adhesive shear damping and to illustrate how they can be utilized a specific example shall be briefly discussed. In this case it will be helpful to recall the general concept[11] that in order to maximize damping, it is necessary to maximize the "effective" stress (or strain) on the interface layer. Having high shear stress locally only on an interface, while allowing most of the interface to remain at a relatively low shear stress, does not produce a high "effective" stress; most of the interface must be at reasonably high stress. Stated differently the integral of the product of local stress (or strain) and interface volume must be maximized.

Let us examine first a relatively simple case of optimizing a design for maximizing damping. Consider the circular plate[16] shown in Fig. 8A held so that its edge is always at zero slope but free to move radially as shown. As the plate vibrates with a center amplitude ζ, the periphery or edge of the plate will move radially with an amplitude s. If now viscoelastic material is placed between the edge faces of the plate and its support, as shown in Fig. 8, the radial motion s provides a mechanism for dissipating damping energy through shear in the viscoelastic interface layer. It can be shown that the amplitude s of radial motion increases with the square of plate deflection δ. The equations for the damping energy dissipated by shear in the viscoelastic material in this configuration were recently developed by Mentel.[16] These equations may be expressed in terms of a dimensionless damping energy D_0^v dissipated per cycle by the viscoelastic material. Damping D_0^v may be maximized by optimizing any one of several different design or material parameters. In order to illustrate this concept and its application, let us consider the problem of selecting the most effective viscoelastic layer by maximizing the damping equation. The manner in which the total energy dissipated by a viscoelastic layer varies with its G_1 and G_2 values (as defined in Eq. 6) is shown in Fig. 8. The damping surface shown indicates that the damping energy dissipation changes as follows: (1) it increases with decreasing values of G_1, and (2) first increases with increasing values of G_2, generally reaches a maximum at some optimum value of G_2 (as shown by the hump in the damping surface), and then decreases. The optimum combinations of values for G_1 and G_2 which result in maximum damping energy dissipation for this special case is shown by the curve traced in the $G_1 - G_2$ base plane. This trace shows that the desirable optimum value for G_2 is approximately 9 and this in combination with a value of G_1 less than 10^2 maxi-

[14] Adhesive No. 466, Minnesota Mining and Mfg. Co.

[15] "Rheological Properties of Adhesives Considered for Interface Damping," by J. S. Whittier and B. J. Lazan. Appendix B in Progress Report 57-6 on "Properties of Materials and Parts under Alternating Force," for Contract AF-33(616)-2803, Wright Air Development Center, Dayton, Ohio, December, 1957.

[16] "Damping Energy Dissipated by Interfaces in Beam and Plate Supports and in Sandwich Cores," by T. J. Mentel. WADC TR 58-547, ASTIA Doc. 206,667, December, 1958.

mizes the damping.

It should be emphasized that scales of Fig. 8 are logarithmic and that the differences in damping energy dissipation for different combinations of G_1 and G_2 may be very large. For example, for the combination of G_1 equal to 10^6 and G_2 equal to 1, the damping energy dissipation is less than 10^{-17}; whereas under the optimum condition of G_2 equal to 9 and G_1 equal to 10, the damping energy dissipation is more than 10^{-9}. The ratio of these two values is 10^8. The potential gain which can be realized through the optimization process is therefore large. This comparison does not, of course, consider material damping in the panel itself.

It would be instructive for general guidance to compare for this case the energy dissipated by the hysteretic effects in the structural material of the panel with the viscoelastic interface damping discussed above. This is done in Fig. 9 which shows the damping as a function of panel dimensions h/a and panel center deflection δ. The damping of the viscoelastic interface, optimized as discussed previously and shown in Fig. 8, is indicated by the D_0^v plane. The energy which can be dissipated through hysteretic damping in the structural material itself is shown on a similar basis by the D_0^m plane. Observe that at low values of h/a for the panel the viscoelastic interface damping is large compared to the material damping. The two damping values are approximately equal, for instance, at h/a of 10^{-2} and δ of 10^{-4}, point A on the intersection line AB. For large values of h/a, and particularly for small values of δ, the material damping is larger than the viscoelastic damping.

The total damping of the panel is, of course, the sum of the two contributions. The differences between the two in some regions of the h/a and δ parameters is very large. (Observe logarithmic scales.) It is important, before deciding on how to increase the total damping in a system, to know which of the two contributors offers the greater potential. For example, in those regions where viscoelastic damping dominates, relatively little can be gained in total damping by material substitution, heat-treatment, or mechanical processing for higher damping. If, for example, the material damping is only 1% of the total, a 1000% increase in material damping will increase the total damping by only 10%. Therefore, before total lamping can be increased significantly it is first necessary to decide, by analyses such as indicated above, wherein the greater potential gain lies. Then a decision can be made on whether to concentrate on improving the materials or the design.

An analytical approach such as outlined above also enables studying the effects of changes in design and material parameters other than those considered.[16] Curves and surfaces similar to those shown in Figs. 8 and 9 may be constructed to clarify the significance of many different design parameters.

"Design Optimization" and "Surface Addition" Approaches for Increasing Damping — The approach discussed above might be identified as the "design optimization" approach. Its object is to suggest the proper proportions and materials for a configuration which must have high damping. In practice it is usually difficult to change an existing design, yet there is often room to add a surface treatment to

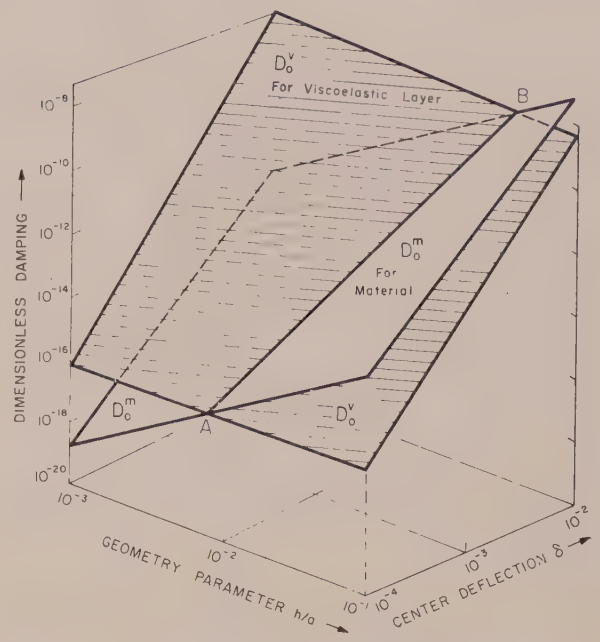


Fig. 9 — Comparison of adhesive interface damping and material hysteretic damping for circular plate

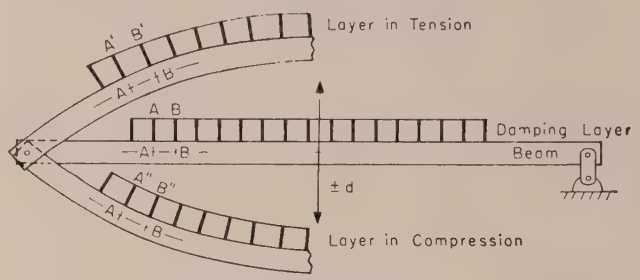


Fig. 10 — Viscoelastic coating which dissipates hysteretic damping energy under cyclic direct stress (tension-compression)

increase the damping of a member which otherwise cannot be changed radically. Examples of such a "surface addition" approach are damping coatings, damping tapes, spacing layers, sandwich additions, and other types of layered construction. Various types of surface additions have been tried in the past. Three of the more important types are described below.

One of the earliest approaches for increasing the damping of panels is to add a thick layer of viscoelastic or similar type of material to a panel (identified by heavy cross-hatching in Fig. 10). As the panel is subjected to cyclic bending deflections $\pm d$, cyclic direct strain (tension-compression) is developed in the viscoelastic layer, as shown in Fig. 10 by the difference in the lengths of line $A'B'$ ($> AB$) and $A''B''$ ($< AB$). Although reasonably satisfactory in some types of panel noise problems, this approach for increasing damping is generally ineffective, since the cyclic strain in the viscoelastic coating is usually small and only a small percentage of the damping capacity of the viscoelastic layer is utilized.

A second "conventional" surface addition approach for increasing damping involves the addi-

tion of so-called "sound damping tape." This is a configuration shown schematically in Fig. 11, consisting of a thin layer *A* of viscoelastic adhesive (thickness exaggerated in figure for clarity) and a backing band or tape *T*, usually aluminum or some other metal. The mechanism of energy dissipation in this case is one of shear in the viscoelastic layer, as shown schematically by the distortions in the cross-hatch bands. As the beam flexes, the viscoelastic layer, located well above the neutral axis and restrained by the backing tape, receives a cyclic shear, and dissipates energy. However, as in the case of the viscoelastic coating shown in Fig. 10, only a very small percentage of the damping which the viscoelastic layer is capable of dissipating can generally be realized. For most thin beams or panels the distance between the neutral axis of bending and the plane of the viscoelastic adhesive is so small that the cyclic shear strain imposed on the adhesive is well below its limits; thus, the energy dissipation is also well below the capacity of the adhesive.

An improvement on the damping tape method shown in Fig. 11 involves a spacing layer,[12, 17] shown in Fig. 12A. This spacer locates the viscoelastic layer further from the neutral axis and increases the cyclic shear associated with a given cyclic flexing. Theoretically, the damping energy which can be dissipated by the adhesive shear mechanism can be greatly increased by such a spacing layer; however, there are some practical limitations. In order to be effective, the spacing layer must be reasonably thick; and most applications require a low density material to avoid excessive weight and cost. Unfortunately, low density spacing layers, or core materials, have relatively low shear modulus. Thus, as illustrated in Fig. 12B much of the cyclic shear which would otherwise be transmitted to the viscoelastic adhesive may be lost as shear strain in the core material. For a rigid spacing layer (Fig. 12A) all the cyclic shear strain γ_o is felt by the adhesive, whereas for a low density spacing layers having relatively low shear modulus (Fig. 12B), the available total cyclic shear γ_o is partially dissipated in the core (part γ_c) leaving only part γ_a for the adhesive layer. As in the previous cases, the adhesive will be subjected to relatively small cyclic shear and the full potential of the viscoelastic adhesive is again not realized.

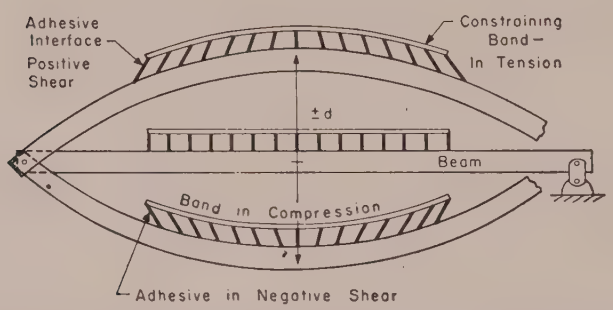


Fig. 11 — "Damping tape" which dissipates hysteretic damping energy under cyclic shear strain in adhesive between beam and constraining tape

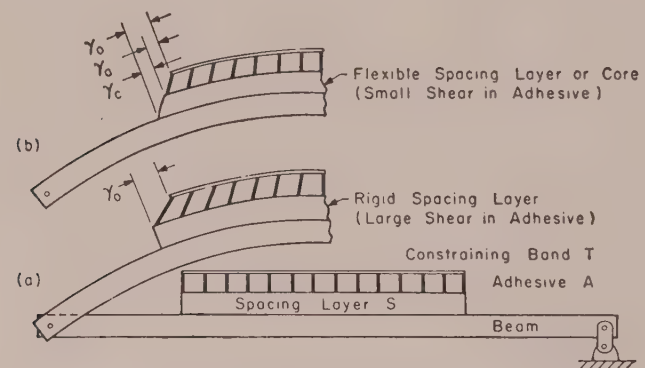


Fig. 12 — Damping tape on spacing layer

A spacing layer approach can also be used to increase the distance between a damping layer, such as shown in Fig. 10, and the neutral axis of bending, increasing the direct (tension-compression) cyclic strain. However, the same difficulty discussed above is again encountered; low shear modulus of low density spacers will absorb much of the available shear strain.

Other surface addition approaches are currently under study.[18] As in the case of the configuration optimization approach, it is important in the surface addition approach to optimize the design for maximum damping carefully. As an illustrative problem, consider the solid beam *AB* shown in Fig. 13 to which lugs have been added as shown. This simulates the spacing layer approach illustrated in Fig. 11A in that the viscoelastic shear region is moved a greater distance ρ from the neutral axis *NA* of the beam and the cyclic shear strain in the viscoelastic layer is correspondingly increased. Equations which have been developed for this case[17] show that, as in the previous case, it is now possible to optimize any one of several design and material parameters and by so doing to maximize the damping. However, for illustrative purposes we consider just two of the design parameters: the thickness *b*, and the property G_1/G_2 of the viscoelastic layer. For the particular beam considered the total damping energy as a function of these two parameters may be plotted as shown in Fig. 13. It is observed from this figure that the optimum thickness for a variety of viscoelastic materials (various values of G_1/G_2) is approximately 5 mills. Furthermore, the lower ratio G_1/G_2 the higher the damping energy dissipated at the viscoelastic layer. As the thickness decreases below two mills the damping decreases very rapidly, but as the thickness increases above 5 mills the corresponding decrease in damping is relatively small. Observations such as these help indicate how critically the design must be optimized.

As in the previous case, it is desirable to compare the damping associated with the hysteretic effects in the beam material itself with damping associated with the viscoelastic layer. Such a comparison is made in Fig. 14. The damping of the viscoelastic layer as a function of two parameters (representing the material and the geometry factors as defined in Fig. 13) optimized to produce maximum damping is shown by the curved surface.

[17] "Effect of Configurational Additions Using Visco-elastic Interfaces on Damping of Cantilever Beam," by J. S. Whittier. WADC TR 58-568, ASTIA Doc. 214,381, May, 1959.

[18] "Resonant Vibration and Acoustic Fatigue Considering Material and Interface Damping." Status Report 58-4 on WADC Contract AF 33(616)-5426, by B. J. Lazan, University of Minnesota, December, 1958.

For comparison purposes the hysteretic damping of the beam material, itself stressed to its fatigue limit, is shown as a horizontal plane having an elevation of 0.0037. The intersection of this horizontal plane representing material damping, and the curved surface representing viscoelastic damping, curve *ABC*, separates regions in which each mechanism dominates. Observe that for very low values of the material and geometry parameters, the material damping may exceed viscoelastic damping; and that for intermediate and high values of these two parameters, the viscoelastic interface damping is much larger than the material damping. At the best combination of the two parameters shown in the figure the viscoelastic damping is approximately 20 times as large as the material damping.

Other cases illustrating design factors important for high viscoelastic damping are discussed in Mentel[16] and Lazan.[18]

Conclusions

1. In the past either the natural frequency of a system or its exciting frequency could generally be adjusted to avoid resonant conditions. However, this is no longer possible in many types of aerospace vehicles. In particular, many types of random excitation, especially those of acoustical origin, cover such a wide frequency band that resonant vibrations can no longer be avoided.

2. The design concepts of former years are inadequate to deal with the fatigue problem associated with resonance amplification. Although "beefing-up" a structure for increased strength and stiffness has "fixed" resonance fatigue difficulties in some cases, this approach cannot be considered a long term engineering solution. Not only is the cost and weight penalty large in such an approach, but also it is totally inadequate to meet the problems encountered in many of the newer types of aero-space vehicles. Thus, the greater utilization of structural damping as an engineering property to control resonance is a necessity. Looking to the future it will still be necessary, of course, to stiffen panels and other configuration by using sandwich and other similar buildups. However, if the maximization of damping is also considered to be a design objective, in addition to the stiffness increase, then a much larger gain in resonant fatigue strength can be realized than using either criterion alone to the exclusion of the other.

3. Whereas in the past the damping of a structural assembly was often increased by the addition of separate energy absorption units (dashpots and the like) or energy transfer devices (dynamic absorbers), this approach no longer provides an engineering solution for many of the newer types of configurations and resonance conditions. Looking to the future the design concept of *optimizing a configuration for maximum inherent structural damping* (damping built into the structure itself) must be emphasized.

4. The experimental and analytical knowledge on the major sources of *inherent structural damping* (hysteresis in structural materials, interface slip, and interface adhesive shear) have now reached a point where application to structural design is practicable. Interface slip must, however, be used with caution because of the fretting and corrosion it may produce.

5. The relative importance of hysteresis in the structural materials and interface adhesive shear as damping mechanisms depends on the configuration, how it is optimized, and the force regimes considered. In most cases, however, structural materials are selected and shaped for properties other than large damping; whereas the interface adhesive may be especially selected and the layer thickness optimized for large damping. In such an optimized design, the contribution of interface adhesive shear to the total damping of a structural system can be very significant.

6. Looking to the future the greater utilization of interface adhesive shear as a damping mechanism, both in an optimized configuration and in surface addition, should be encouraged.

Acknowledgments

Much of the work reported in this paper was performed under the sponsorship of Wright Air Development Center, Wright-Patterson Air Force Base, Ohio, under the general liaison of Walter J. Trapp of the Materials Laboratory.

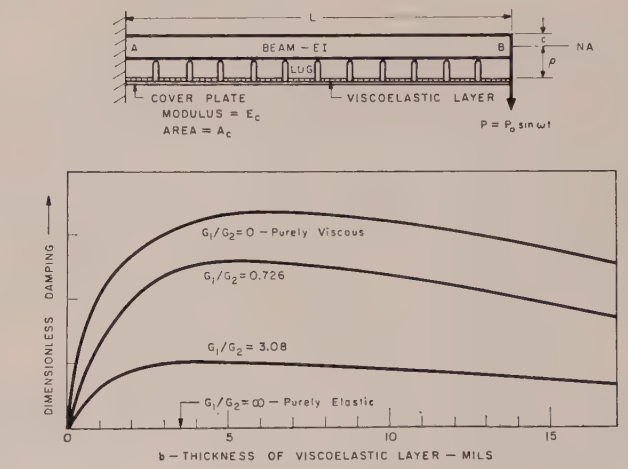


Fig. 13 — Effect of thickness of adhesive layer on damping due to surface addition

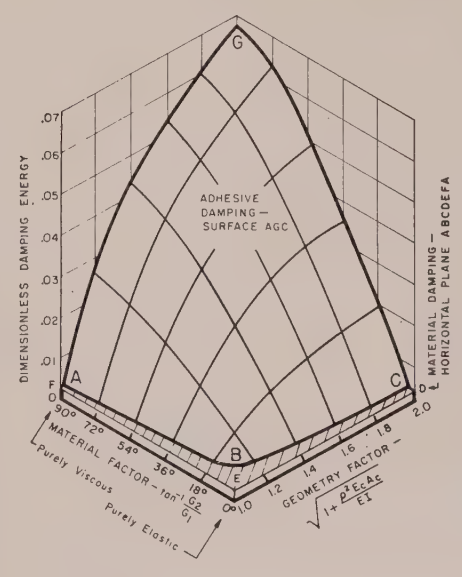


Fig. 14 — Comparison of optimized adhesive damping with material damping of beam

THIS PAPER explains the meaning of static, kinetic, and dynamic unbalance and how these different forms can be recognized or avoided in the maintenance work performed on commercial engines and vehicles.

Critical speeds, resonant frequencies, their effects and means for avoidance or changing, as well as other vibration problems are discussed with the layman in mind. All examples are taken from actual case histories.*

Dynamic of rebuilt

Ralph T. Buscarello

Stewart-Warner Corp.

PRECISION BALANCING of the rebuilt automotive engine is not new, especially among race car drivers, hot rod enthusiasts, and sport car owners. These people have for years been interested in extra performance obtained through a precision balanced engine. Their main interest usually has been to win a race or to get better control of the car under competition, but the "side results" have been of sufficient value to be of interest to users of commercial vehicles.

Commercial and Precision Balancing

It is known that automotive engines are balanced at the factory to a commercial tolerance. However, mass production methods allow a tolerance that is much greater than that of the precision balanced engine. No formal tests have been made to determine what closing down this tolerance has meant, but there are now so many reports from those who have had their engines precision balanced that a picture has evolved. For example, bearing life on higher speed engines has increased as much as 300%. For commercial truck engines operating at lower speeds, reports have indicated an increase in engine life from 25 to 100% between tear downs. Increased horsepower on the average has been approximately 10% with the same percentage gain in acceleration. Within the last few years, many truck engines have been precision balanced to the same tolerance as the finest racing engines and have enjoyed the results. The trend has started but generally in localized areas. In these cases the maintenance foreman usually was a hot rodder as a kid, and brought his experience in balancing with him when he reached his present position.

Balancing Pistons and Connecting Rods

For a complete precision engine balance, each piston must be brought down to the weight of the lightest piston. First, the weight of each piston is recorded on a card and the weight difference between the heaviest to the lightest piston noted. Precision balancing of pistons require a tolerance of + ½ gram, minus zero. Experience indicates that truck pistons usually have a commercial tolerance of 6–15 grams, whereas passenger-car pistons vary 6–8 grams between the lightest and heaviest units.

Each piston is lightened by removing metal from inside the skirt. The connecting rods must also be balanced, not only as total rod, but end for end. That is, all of the rotating ends, as well as the reciprocating or piston pin ends, should be brought to within ½ gram of each other. Each end must be balanced separately, for the reciprocating end has an entirely different effect on the crankshaft compared to the pure rotating force of the large end. Typical difference in weight of new connecting rods of commercial engines is usually from 8 to 20 grams. With rebuilt connecting rods, the difference between the lightest and heaviest varies as much as 30–40 grams. In many cases, they are so far out of balance that not enough material can be removed and another set has to be obtained. Metal is removed from the balance pads usually found at each end of the connecting rod.

Bob Weights

In the case of V crankshafts, the effective weight of the rod and piston assembly affects the dynamic balance of the crankshaft itself, even if the crankshaft was not modified during the rebuilding process. For this reason, the operator has to record and weigh a complete set of rings, piston pin, rod insert bearings, and even the pin locks. A bob weight takes the place of the rod and piston assembly during the balancing of the shaft on a balancing machine. Each bob weight is brought to the exact weight computed and then clamped onto each individual throw. The simple formula for this bob weight is 100% of the rotating weight about each throw plus 50% of the reciprocating weight affecting that throw. After the bob weights are mounted, the V shaft can be balanced in the same manner as an in-line crankshaft. All 4-, 6, and 8-cyl in-line crankshafts can be balanced without bob weights. Another crankshaft design requiring bob weights is found in the 71 series GMC diesels. In this case, however, the bob weight is 100% of

* Paper presented at SAE National Transportation Meeting, Chicago, Oct. 26, 1959.

Balancing

engines and driveline components

the rotating weight only, with the reciprocating weight taken care of by rotating counterweights apart from the crankshaft.

Balancing Rotating Parts

For balancing rebuilt engines to precision tolerances, commercially available job-type dynamic balancing machines are necessary because the same balancer must be capable of balancing the largest crankshaft as well as the smallest pulley. On the balancer, the crankshaft is usually balanced first, after which the flywheel is mounted on the balanced shaft and balanced to the same precision tolerance. Next the clutch is mounted in place and balanced in the same manner, followed by the torsional dampener at the other end. With proper mandrils, the flywheel and clutch assemblies can be balanced individually, apart from the crankshaft itself, thus enabling easier handling.

Balance Tolerance for Rotating Parts

Tolerances for dynamic balancing vary with the engines and companies involved. However, the average commercial engine crankshaft seems to have a balancing tolerance of approximately 1–2 oz-in. Precision balancing requires a balancing tolerance of less than 0.1 oz-in. Decreasing the unbalance in the rotating part decreases the centrifugal force used to vibrate the engine. This vibrating force utilizes energy that should go into useful work rather than into wearing out parts. Higher speeds of modern engines are new super highway programs have made vibration a real problem — a truck that used to travel at high speed for only short intervals of time, can now maintain high speed for hours on end. Effects of vibration at high speed for a short time are not too noticeable on comfort and fatigue, but the same vibration for relatively long periods of time has caused the driver to complain and request that something be done about it. In many cases, balancing tolerances perfectly acceptable just a short time ago are not acceptable under the high-speed and long-time conditions found today.

Unbalance Centrifugal Force

The formula for centrifugal force is:

$$F = mr\,\omega^2 \qquad (1)$$

where:

F = Centrifugal force due to unbalance, lb
m = Mass of unbalance, lb
r = Radius at which unbalanced mass is acting, ft
ω = Proportional to rpm

Doubling the mass of unbalance will double the centrifugal force $\omega^2 = \left[\frac{2\pi(\text{rpm})}{60}\right]^2$ according to the formula. The same will apply in doubling the radius; but doubling the speed will increase the force by the square of the speed, resulting in four times the amount of force. By quadrupling the rpm, the force would not be four times as great but sixteen times greater. It is evident that in an unbalanced system the speed could be increased until the centrifugal forces become excessive and the vibration is well beyond the accepted tolerance.

Because of the varied terminology used in balancing, it is best to give a definition of static, kinetic, and dynamic balancing. An understanding of the principles involved will enable potential sources of each type of unbalance to be located, thereby eliminating the resulting vibration.

In *static* unbalance, only gravity or weight force is involved. For example, assume a rotor which is perfectly balanced and, therefore, does not vibrate regardless of the speed of rotation. If this rotor is placed on frictionless rollers or "knife edges" and a

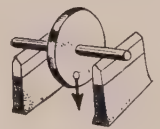


Fig. 1

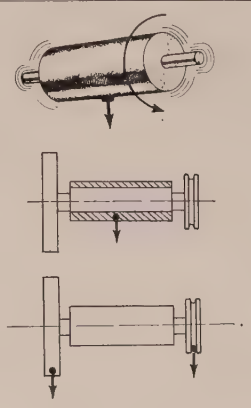


Fig. 2

weight is applied on the rim at the center of gravity line (between the two ends), the weighted portion is acted upon by gravitational force and immediately rolls to the bottom. This condition is known as static unbalance. Balancing this rotor, without rotation, may be accomplished provided the rotor is a narrow, inflexible type as illustrated in Fig. 1.

Kinetic unbalance is "static" unbalance in rotating motion (Fig. 2). Kinetic unbalance occurs when the rotor spins and the unbalance weight sets up a centrifugal force which causes deflection of the rotor in the same direction as the unbalance force. If the rotor has any flexibility at all, it will deflect or bend in the direction of the original static unbalance weight, thereby increasing the amount of distortion, that is, radius r. This deflection of the rotor causes still greater centrifugal force, resulting in still greater deflection. Static or kinetic unbalance in a rebuilt engine can be caused by a crankshaft ground off-center equally at both ends, causing one side to be heavier than the other. It can also be caused by unbalance in the flywheel and clutch assembly in the same direction as the unbalance in the torsional dampener. (See Fig. 3.)

For narrow objects that are relatively rigid, static unbalance and kinetic unbalance can be treated alike; whereas with longer, more flexible rotors, kinetic unbalance alone governs correction.

Some balancer manufacturers refer to "dynamic" balancing as any type of rotating balance; but technically speaking, dynamic unbalance is strictly a "couple" or twist caused by two equal forces in two separate planes, 180 deg apart opposing each other. (See Fig. 3.)

Assume the rotor in Fig. 3 is placed on frictionless rollers. As the weights (forces) are 180 deg apart and in the same plane through the longitudinal centerline of the shaft, this rotor is balanced statically and kinetically. If this same rotor is revolved at an appreciable speed, each weight would then cause a centrifugal force, which acting in a separate plane of rotation would cause a rocking motion and tend to twist the rotor as indicated by the curved arrow. This type of unbalance is pure dynamic unbalance. This can be simulated by placing a pencil on a table and pushing with one finger in one direction, as you push with another finger in another plane in the opposite direction. The pencil will tend to rotate "end for end." This end-for-end action is restricted in a rotating rotor by the bearings that hold the rotor in one plane — however, this tendency to want to turn end-for-end is still present. It is this "twist" force that causes the rotor to vibrate. Rarely does pure kinetic or dynamic unbalance occur in rotors. For example, assume that a weight is placed on a perfectly balanced rotor, to one side of the center of gravity line (Fig. 4). This weight would cause the rotor to turn until the weight was on the bottom of the rotor (static unbalance). However, rotate this rotor at an appreciable speed and the centrifugal force, due to the unbalance, will result in a twisting action caused by its trying to rotate about its center of gravity. The extent to which this tends to twist or form a "couple" is the dynamic unbalance. This combination of kinetic and dynamic unbalance is usually referred to as "dynamic" unbalance; however, for clarity the remainder of this paper will refer to dynamic unbalance as pure dynamic couple or twist, separate from kinetic force.

In the rebuilt engine, dynamic unbalance may be caused by the crankshaft being offset a few thousandths at each end in opposite directions relative to each other, or by the unbalance in the flywheel and clutch assembly end acting in the opposite direction to the unbalance in the torsional dampener. However, the main source of dynamic unbalance in the V engine is from the effect of the different weights of the rod and piston assembles relative to the original weights to which the shaft was balanced. Very often no care is placed in selecting the rod and piston assemblies according to weight. Pistons taken from a different production run can often vary enough in weight to cause considerable dynamic unbalance on the crankshaft. This would cause an entirely different bob weight with its resulting change in dynamic unbalance. Even if the crankshaft were not touched in the rebuilding process, it would have to be rebalanced in order to eliminate a "vibrator."

Fig. 4 is a combination of kinetic and dynamic unbalance. Fig. 5 breaks this single force into its component dynamic couple and kinetic force. This condition can be caused on a well-balanced assembly by placing a balanced flywheel on a balanced crankshaft but allowing gravity to take up the play between the crankshaft flange and the flywheel, in one direction. (See Fig. 6.) The mass of the off-center flywheel results in a kinetic unbalance within the flywheel and a resulting combination of kinetic and dynamic unbalance relative to the whole assembly. Some rebuilders are now pinning the clutch assembly into place before precision balancing, so that the mechanic can replace it in the same position in which it was balanced. This also forces the mechanic to mount it in the same indexed position, eliminating the unbalance that results when it is indexed around to another angle. Another cause of dynamic unbalance in an otherwise balanced assembly is the cocking of the flywheel due to dirt or burrs between the face of the crankshaft flange and the flywheel mounting surfaces. Just a few thousandths of an inch of dirt to one side can cause enough cocking of the flywheel and clutch assembly to result in a dynamic unbalance that is well beyond the allowable tolerance. Therefore, precision balancing of engines not only requires care during the balancing procedure but also in the as-

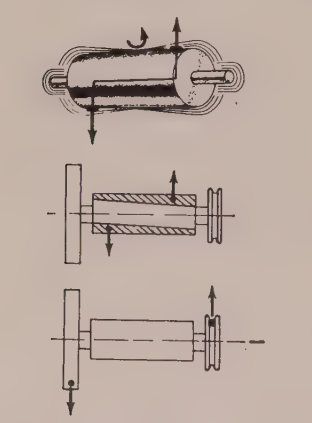


Fig. 3

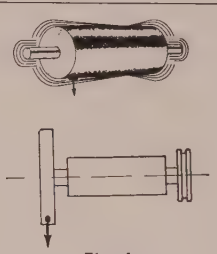


Fig. 4

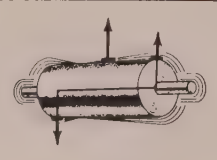


Fig. 5

sembly operation. Dynamic unbalance also occurs in the fan assembly when one blade is pushed out axially relative to its diametrically opposite blade. Usually, bringing this back into dynamic balance is a matter of bending blades to the proper position.

Balancing Drivelines

Another problem has recently arisen with the increasing speed of long rotors and flexible shafts, such as drivelines. Fig. 7 shows a long driveline with an unbalance weight near its axial center. For convenience, the operator usually chooses the end planes to make his corrections. The large distance between the actual unbalanced weight and the counterbalancing corrections in the end planes causes a deflection of the rotor at operating speed. The rotor is deflected in the same direction as the original unbalanced weight, causing further unbalance; however the correction in the end planes causes less deflection, resulting in an insufficient amount of correction. If the shaft were short and rigid, this deflection would be negligible and the rotor would run smoothly at operating conditions. Better results could be obtained by dividing the long rotor into three parts, such as in Fig. 8. Static and dynamic balance corrections would then be made in the two inside planes. Note that the distance between any unbalanced weight and the counterbalancing weights is never very great relative to the rotor diameter. The deflections at operating speed would be negligible and, therefore, the precision balance obtained on the balancing machine would also apply at operating speed.

Another way to handle vibration problems in drivelines is to balance these units after they are mounted in the vehicle using a commercially available portable balancer. The ease in which weights can be added to drivelines even after mounting may make the portable balancing method the most practical, as it takes into consideration all of the effects if misalignment as well as actual unbalance. One precaution in balancing drivelines in the field is to not get the shaft too hot with an arc welder, as this would cause the shaft to bend and give an erroneous reading. After the shaft cools off, it would then be still out of balance.

When arc welding is the only method available, some operators tack weld the counterbalancing weights on and then run a bead exactly opposite this position in order to get an approximately equal amount of heat on the other side. For trial runs in portable balancing, aviation-type hose clamps are usually used to hold washers in place. The washers are then weighed and a permanent counterbalancing weight substituted. On some types of drivelines, it is best to not use a hammer to knock off the excess slag, as this can cause a slightly repositioned shaft with its resulting unbalance. The light slag is usually negligible in weight and, therefore, should be left on to come off by ordinary abrasion during operation.

Balance Tolerance for "Flexible" Rotors

Another problem caused by the deflection is exhibited in the case of a rotor mounted in a cantilevered position on the end of a shaft rather than mounted between bearings. In this case the kinetic balance tolerance would have to be 5–10 times closer for the cantilevered rotor than for the rotor mounted between bearings. Fig. 9 shows a small amount of allowable unbalance with a pump impeller mounted between bearings. The same pump impeller and unbalance condition mounted as in Fig. 10 may cause excessive vibration. Even the smallest amount of unbalance may result in the shaft deflecting at operating speed, decentering the mass of the rotor in the same direction as the original unbalance and thereby adding to it. This new unbalanced condition would cause further deflection, further unbalance, and so on — until equilibrium is obtained with an excessive vibration. Whenever a part has been balanced on a balancing machine to the usual tolerances and still vibrates or causes further vibration, care should be taken to see whether the method of mounting the part is the contributing factor to the excess of vibration. By decreasing the balancing tolerance, the trouble can be removed.

Effects of Critical Speeds or Resonance

The vibratory effect of a part that is in good dynamic balance as well as kinetic balance may still be magnified through the structure of the machine. Every driver of a motor vehicle should be familiar with the fact that his vehicle seems to have more vibration at certain speeds than at others. According to our knowledge of and in accordance with the formula for centrifugal force, the vibration should increase steadily as the rpm increases. Instead the vibration increases with the rpm, suddenly increases at a much higher rate through the "critical speed," and then apparently smooths out as the rpm goes beyond this critical speed. What has happened is that the frequency of the vibration due to unbalance is the same as the natural frequency of the framework or of other supporting members. To visualize this phenomenon, consider a simple flat spring with a weight mounted at one end as, in Fig. 11. Deflect the spring by pulling down on the

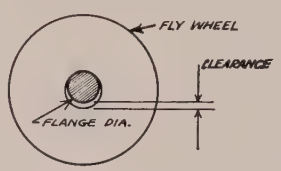


Fig. 6

Fig. 7

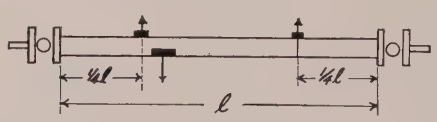


Fig. 8

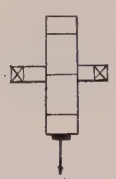


Fig. 9

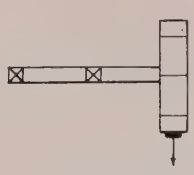


Fig. 10

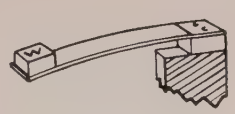


Fig. 11

weight and then let the weight go so that the spring will oscillate up and down. The weight will vibrate at a definite frequency. Once a spring is deflected in this way, it will have a natural frequency of vibration that will remain the same until it comes to a complete stop. Assume now that you can give a slight impulse to the spring in either direction during its upward travel or during its downward travel. If the frequency of the impulse is not the same as the natural frequency of the spring, sometimes the weight would be on its way up when the impulse given would be in the downward direction. Sometimes the amplitude of vibration would be decreased by the impulse, whereas at other times the impulse would cause the amplitude of vibration to become greater. The resultant effect would be an irregular vibration that does not build up.

Now consider the effect when the impulse given to the spring both in the upward and downward direction, coincides with the natural frequency of the spring. Assume that the spring and weight are at rest and the first impulse at a given frequency is now transmitted to the weight. Because the weight is deflected once (by the impulse) and allowed to vibrate back again, the spring would vibrate at its natural frequency even if no other impulses were given. But there is another impulse given, at the same frequency of vibration as the natural frequency. Just as the spring reaches bottom dead center, the impulse would give the weight a push in the upward direction and as the weight reaches the top of its travel, the impulse would give a push in the downward direction giving added amplitude of vibration. This process would continue to go on giving a greater and greater amplitude or distance traveled by the spring. The frequency of vibration, however, would remain the same. If this process were allowed to go on continuously, the spring would either break or reach an equilibrium amplitude of vibration. Experience has shown that resonance magnifies vibration (undamped) 100–200 times the amplitude normally attained by the same amount of unbalance operating at a nonresonant speed.

This same process can be explained in a simple manner by visualizing a man on a diving board, which in effect is a flat spring deflecting at a fulcrum point. When the diver jumps up and down on the end of the board, he very quickly observes that when he synchronizes his efforts with the motion of the board, it takes very little exertion to keep the board in motion and make it bend at its maximum amplitude. Every machine part or assembly has a critical or natural frequency. Design engineers try to keep the natural frequency higher than the highest rpm of any rotating part, or sometimes considerably lower than normal operating speeds. Sometimes the frequency of vibration may be brought up to resonance by the individual sections of fan or turbine blades passing a point where they lose part of their load, such as at an outlet. For example, a machine's frame may have a natural frequency of 8100 cpm. The maximum fan speed may be only 2000 rpm but its four blades passing an outlet cause a vibration of 8000 cpm. This is close enough to cause the frame to get some of the magnification effects of resonance.

An interesting case involving critical speeds or the resonant frequency occurred on a truck engine mount. The unit vibrated excessively due to a resonant condition and all attempts to remove this vibration met with virtually no success. The final solution was to weld some $\frac{5}{8}$ in. diameter steel rods along the underside of the channel beam cross-piece supporting the engine. Two $\frac{5}{8}$-in. rods were placed inside each leg of the channel. It was found that solidly welding the rods to the channel did not improve the condition very much, but tack welding the rods in place at about 8-in. intervals did — with each alternating tack occurring at the side of the rod to the channel leg and the other tack welded to the top of the channel beam. It was felt that the rods would give additional stiffness to the beam, thereby bringing its natural frequency to above operating speed. However, the real solution was due to the fact that the rods were free to vibrate apart from the channel beam itself, because the alternate tack welds were spread enough apart so the vibrating beam "bounced" against the added rods. This minute "bouncing" results in friction which has a drastic effect in reducing vibration due to resonance. The friction damping allowed the vehicle to be acceptable by the user.

Another case of vibration removal through handling the resonance problem occurred on a vehicle that resonated at the typical operating speed desired by the driver on the open road. Precision balancing the driveline brought no results. Finally, all of the bolts on the frame were tightened and two cross member I-beams were welded in place to stiffen the vehicles framework. (To change the resonant or critical frequency of a machine or vehicle, it is necessary to change the stiffness. A long, flexible part would have a lower critical speed than a short rigid part. Often the critical speed of a part is raised by making it more rigid in order to get a frequency higher than that of the offending vibration. Sometimes it is impossible because of other design limitations to increase the size of the part to stiffen it. In some cases the only solution may be to make a part less rigid or by adding weight at the right place to get greater deflection, thereby lowering the natural frequency to a point below the normal operating range.) After this stiffening, the natural frequency of the frame was beyond the operating range and the vehicle produced excellent results.

Use of Portable Balancer

Throughout this paper, kinetic and dynamic balancing was assumed to be performed on a cradle balancer where the work is taken to the balancing machine. Often, the fastest solution is to use a portable balancer, which not only allows the balancing of the parts in the assembled condition, but also allows the operator to determine the part causing the offending vibration. Commercially available now are portable balancers which indicate the frequency or rpm of the offending vibration and its amplitude. With this information an operator can measure the vibration at each end of the crankshaft, at the fan assembly, and at the driveline bearings. By noting the frequency and amount of the worst vibration, he can then pin-point the main source of trouble. It is not possible at the present moment to determine whether the unbalance is at the flange end of the crankshaft itself or in the flywheel mounted to that flange, as the resulting fre-

quency in both cases would be exactly the same. It does, however, tell the operator whether to work on the flywheel and clutch assembly end or the fan, driveline, and torsional dampener end. This "zeros-in" the area where the offensive vibration emanates, eliminating a large amount of guesswork which can result in long hours interchanging parts.

Another use of a portable balancer is to determine whether the offensive vibration is beyond a predetermined acceptable level or not. Quite often a driver may complain of a vibration on one day that will be perfectly acceptable when he is in a better mood. Having the foreman feel the vibration with his fingers and then giving his opinion, plus the operator doing the same, can lead to arguments with no real solution whereas the unbalance meter reading is pretty hard to beat.

Which Balancing Method?

For maintenance work the advantages of a portable balancer are:

1. Sometimes the operator can balance the part in less time than it would take to remove the part.
2. It is possible to locate that part which requires balancing among many rotating parts.
3. Whether a part is vibrating beyond allowable limits is determined before attempting to balance.
4. Sources of vibration other than the rotor can be balanced, such as the drive belt, misaligned shafts, and the like.
5. The final balance is obtained in the final assembled condition, which would be important if the part could not be accurately centered.

The advantage of a job-type maintenance cradle balancer are:

1. Balacing can be performed to a much closer tolerance (even closer than that which would cause no vibration) than with a portable balancer, thereby giving less load at the bearings.
2. When an engine has to be dismantled for other repairs such as replacing bearings, rotating parts can be balanced while other work is performed, enabling operation to be resumed as soon as the engine is assembled.
3. Balancing a rotor on a cradle balancer is usually four to six times faster than balancing the same rotor after assembly.
4. Balancing corrections can be made without affecting other machine parts with grinder grit, weld splatter, or drill chips.
5. Balancing a rotor on a cradle balancer gives greater accessibility in making balancing corrections. Sometimes it is almost impossible to get at the rotor to make corrections when balancing with a portable balancer.
6. It is more difficult to train an operator for portable balancing than for cradle balancing, due to the infinite number of outside conditions affecting the vibration of assembled machinery.

Precision kinetic and dynamic balancing used to be reserved for the manufacturers of equipment. However, with the job-type balancing machines developed for maintenance applications, the advantages of balancing are now available for the users of vehicles and engines. Downtime to replace bearings or to repair a part that has failed by progressive fracture (often referred to by mechanics as "crystallization") can be so expensive that decreasing the number of shutdowns easily justifies the expense of balancing to remove vibration. The minimum return from precision balancing (over and above commercial balancing tolerances) reported by truckers who have kept accurate records, has been 25% increase in engine life between tear downs. Reported increase of 100% is as common. The figure varies, of course, with the operating speeds of the engines involved. Seal life has been increased, as well as fuel consumption decreased. One driver reports "on certain hills where I had to change into a lower gear on each trip, I now can make these hills without changing gears." Another reports "since the engine has been precision balanced, I now make the run in twenty minutes less time." Decreased fatigue is probably the largest one item noticeably by the driver, resulting in need for fewer rest stops as well as a higher safety factor that cannot easily be measured in dollars. One operator reports the unexpected result of noticeably decreased cab maintenance. Originally going into precision balancing as a means to get a greater percentage of the rebuilding business, some rebuilders added the precision balancing job free in order to eliminate comeback, as well as improve the general quality of their work. However, most of them found that their customers were glad to pay the small additional cost of precision balancing once they were able to determine the difference between a rebuilt engine that was precision balanced and one that was not.

Less noise, less irritation, greater safety, longer life of parts, and all around increased performance of the vehicle are not easily measured in dollar value. But every maintenance foreman, who has the problem of keeping his equipment running and running well, knows the value of eliminating destructive vibration.

ORAL DISCUSSION

Reported by D. B. Albrecht
The Standard Oil Co. (Ohio)

Warren A. Toussig, Burlington Truck Lines: Why in the case of a long driveline do you recommend putting the counterweights only one-quarter of the way in toward the center when balancing statically and dynamically? Wouldn't it be better to put it opposite the excess weight?

Mr. Buscarello: So far, there is really no machine available for determination of that exact spot or spots. Therefore, the one-quarter distance has been determined by experience to be the best arbitrary point for elimination or reduction of deflection. On the average, it works out best. This has been confirmed by experience with long shafts or rolls in the paper industry .

S. Willard Lord, Pennsylvania State University: Can Straight shafts are one thing, but in U-joints there is the balancing be used when looking at universal joints? problem of producing the same angularity.

Mr. Buscarello: Misalignment in such parts can be found with the use of portable equipment, although it is harder and takes more experience. It has been found that in many cases the frequency of the vibration is a function of the speed of the moving part; that is, it will vibrate once each revolution or twice each revolution. Therefore, we have been able to pin point particular problems by the stroboscopic image on the equipment. I should like to caution you, however, with the experience of one Denver shop. They found that only about 50% of the vibration problems were caused by unbalance, so the customer should be in a position with other necessary equipment to take preliminary steps of checking before balancing.

RADIATION FROM FLAMES in gas turbine combustors

R. M. Schirmer and L. A. McReynolds
Phillips Petroleum Co

J. A. Daley Naval Air Material Center

THE FUEL QUALITY requirements of the aircraft gas turbine engine have been a source of controversy from the time of its development almost 20 years ago. It is now generally conceded that a jet fuel should be a distillate, free from such foreign materials as water and dirt, and sufficiently refined to insure the removal of objectionable corrosive, gummy, and malodorous constituents.[1] Present military and commercial jet fuel specifications endeavor to define such high-quality distillates. However, it is significant that tolerance for a range in jet fuel properties can be built into the engine.[2] Because of this, there is a need for justification when establishing the allowable specification range in jet fuel physical and chemical properties, which simultaneously requires agreement on the test methods for definition of these properties.

A noteworthy "interpretation" of aircraft gas turbine fuel quality requirements has stressed the need for better control over the basic parameters of burning characteristics. While recognizing the need for a stable and efficient combustion process, it emphasized the inadequacy of present control over the carbon-forming tendencies of jet fuel.[3] Of all the jet fuel properties, those dealing with control over the cleanliness of the combustion process have been debated longest and strongest. Many investigations have been conducted on both the chemical mechanism of the process, as well as the purely physical environmental factors contributing to the formation of smoke and coke in the combustion of hydrocarbon-air mixtures. Much of the available information on combustion cleanliness has been summarized by Schalla and Hibbard,[4] who describe the physical and chemical properties of these materials, the effect of operational and fuel variables on the formation of smoke and coke, and the chemical mechanisms proposed for their formation. In commenting on this situation, Macauley has suggested that the difficulty in evaluating the combustion properties of jet fuels originates from the wide usage of test methods employing simple lamps in spite of questionable accuracy and meaning of the data.[5]

Flame Radiation

General recognition of the major importance of flame radiant energy as a combustion cleanliness problem is undoubtedly the most significant recent development in jet fuel quality control. Pratt & Whitney Division of United Aircraft was first to propose control of flame radiation characteristics by limiting the polycyclic aromatic content (as measured by UV) of jet fuels to 3%, providing the ASTM Smoke Point was less than 25 mm. However, recognizing the limitation of this specification, continued development work was announced on a fuel luminosity lamp.[3] Of particular interest was a correlation of fuel luminosity lamp ratings with the operating temperatures of a full-scale single combustor liner, showing a spread of 500 F in liner temperature resulting from variations in fuel quality alone. Randall and Ritcheske have described the influence of jet fuel composition and physical properties on the operating temperature of gas turbine combustor liners, which again showed a correlation between fuel luminosity lamp ratings and liner temperatures.[6] Subsequently, this fuel luminosity

---

[1] "Critical Survey of Commercial Turbine Fuels," by H. E. Alquist and R. M. Schirmer. Paper presented at SAE National Aeronautic Meeting, New York, April, 1955.

[2] "Must Jets be Pampered?," by H. A. Fremont and E. V. Albert. Paper presented at SAE Annual Meeting, Detroit, January, 1955.

[3] "Relation of Fuels to Aircraft Turbine Engines," by E. A. Droegemueller and R. K. Nelson. Paper presented at SAE National Aeronautic Meeting, New York, April, 1958.

[4] "Basic Considerations in Combustion of Hydrocarbon Fuels with Air," by R. L. Schalla and R. R. Hibbard. Chapter IX, Report 1300, Forty-Third Annual Report of NACA, 1957.

[5] "Cloudy Crystal Ball," by J. B. Macauley. Paper presented at WPRA Annual Meeting, San Antonio, 1959.

[6] "Influence of Fuel Variables on Radiation from Flames in Turbojet Combustor," by N. P. Randall and W. F. Ritcheske. Paper presented at SAE Annual Meeting, Detroit, January, 1959.

THE EFFECT of monocyclic versus polycyclic aromatic components, in JP-5 fuels having the same ASTM Smoke Points, on total flame radiant energy was investigated. The performance of research combustors and a J79 aircraft gas turbine engine single combustor operated at low (atmospheric) pressure showed that variations in aromatic type or content within the present JP-5 specification have no significant effect on flame radiation.

The performance of research combustors and a J57 aircraft gas turbine engine single combustor operated at high (5-15 atm) pressure showed that polycyclic aromatic fuel blends burn with higher flame emissivities than monocyclic aromatic fuel blends of comparable ASTM Smoke Point. Radiant heating of metal parts was shown to be a function of their location in the combustor because quenched combustion products can effectively absorb flame radiation.

A characteristic relationship between total flame radiant energy from a J57 combustor and CRC Luminometer number was obtained with 12 test fuels covering a broad range in burning characteristics. Because of the excellent correlation between fuel ratings obtained with the ASTM Smoke Lamp and the CRC Luminometer, it seems desirable to consider the latter as an alternate for characterizing fuels beyond the limit of the present test method.

The use of low-luminosity fuels gave major reductions in liner temperatures in the J57 combustor. However, such fuels do not insure large reductions in heat transfer to metal parts, as shown by a reduction of only 40 F in the afterburner liner temperature of a J75 aircraft gas turbine engine obtained with "JP-150."*

lamp, now called the CRC Luminometer, was evaluated by CRC for reproducibility and repeatability of fuel ratings,[7] using a test technique which has now been published by ASTM as a proposed test method.[8]

The control of carbonaceous deposits (coke), solids discharged with the exhaust gas (smoke), and radiant heat transfer from the combustion zone (flame radiation) are important factors in aircraft gas turbine combustor design. Control over these combustion cleanliness factors has been sought through both proper design and fuel specification. While little information is available on the detailed factors underlying combustor design, since this is considered a proprietary field, some studies have indicated that combustion cleanliness can be regulated by rigorous control of the fuel's residence time at pyrolytic conditions, that is, by maintaining a subcritical fuel residence time.[9] However, it is significant that there have been no reliable data available on *total* flame radiant energy from aviation gas turbine combustors, and none on the characteristics of their flame spectra. For example, Winter assumed that luminous flame emissivities would be roughly twice the equivalent gray body emissivity of 0.1 he assigned to nonluminous flames for his calculations of convective and radiant heat flow rates to the flame tube adjacent to the burning zone of an aero gas turbine combustion chamber.[10] Berlad and Hibbard used the spectral distribution of energy from luminous and nonluminous propane-air Bunsen-type flames for their calculations of the effect of radiant heating on fuel droplet evaporation rates.[11] Certainly such a lack of basic information about flame emissivity and spectral distribution of energy has not facilitated calculations of radiant heat transfer for proper combustor design. A better understanding of the nature of the processes involved in the exchange of radiant energy among the fuel, flame, combustion products, and combustor components may be useful and necessary for optimum aircraft gas turbine engine design and fuel specification.

Flame Spectroscopy

An investigation of the characteristics of the energy radiated from aircraft gas turbine type combustion processes, using a laboratory scale combustor operated close to atmospheric pressure (40 in. of Hg abs) produced infrared emission spectra shown typically in Fig. 1.[12] The discontinuous radiation of molecular origin, recognized as the familiar water and carbon dioxide vibrational emission bands, was characteristic for both paraffinic and aromatic test fuels. Since the integrated area under the curve represents total flame radiant energy, it can be seen by comparison with the theoretical black body curve for this gas temperature that the equivalent gray body emissivity is very low (about 0.03). The essentially discontinuous spectral distribution of energy is characteristic of most hydrocarbon-air flames at atmospheric pressure, giving low average emissivities regardless of whether they appear luminous or nonluminous to the eye.[13] This necessitates the inclusion of the 1–5 micron spectral region, which can be accomplished with sapphire optics, for reliable measurement of total flame radiant energy.[12] The shape of the theoretical black

---

* Paper presented at SAE Annual Meeting, Detroit, Jan. 11, 1960.

[7] "Evaluation of CRC Luminometer." Report on CRC Project No. CA-16-58, June, 1959.

[8] "Proposed Method of Test for Luminometer Numbers of Aviation Turbine Fuels." Appendix VIII, ASTM Standards on Petroleum Products and Lubricants, October, 1959.

[9] "Film Vaporization Combustor for Gas Turbine Engines," by G. W. Maybach. U. S. Navy Contract No. Nonr 656(13), Task Order No. NR097-341. The Pennsylvania State University, 1959. ASTIA No. AD-209708.

[10] "Heat Transfer Conditions at Flame Tube Walls of Aero Gas Turbine Combustion Chamber," by E. F. Winter. *Fuel*, Vol. 34, October, 1955, pp. 409–428.

[11] "Effect of Radiant Energy on Vaporization and Combustion of Liquid Fuels," by A. L. Berlad and R. R. Hibbard. NACA RM E52I09, 1952.

[12] "Radiation from Laboratory-Scale Jet Combustor Flames," by E. C. Miller, A. E. Blake, R. M. Schirmer, G. D. Kittredge, and E. H. Fromm. Special Report for U. S. Navy Contract NOas 52-132-c. Phillips Petroleum Co. Research Division Report 1526-56R. Condensation of this report was presented at American Chemical Society Meeting, Urbana, May, 1958: "Radiation from Jet Combustor Flames," by R. M. Schirmer and E. C. Miller.

[13] "Diesel Combustion Study by Infrared Emission Spectroscopy," by W. T. Lyn. *Journal* of Institute of Petroleum, Vol. 43, February, 1957, pp. 25–46.

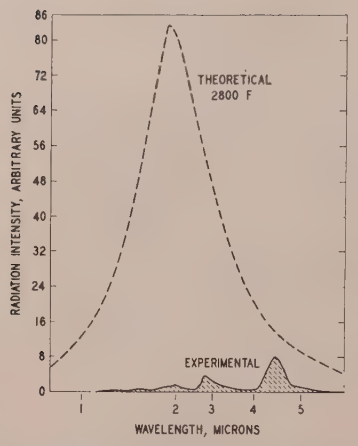


Fig. 1 — Comparison of black body radiant energy with flame emission spectra from 2-in. research combustor at 40 in. of Hg abs

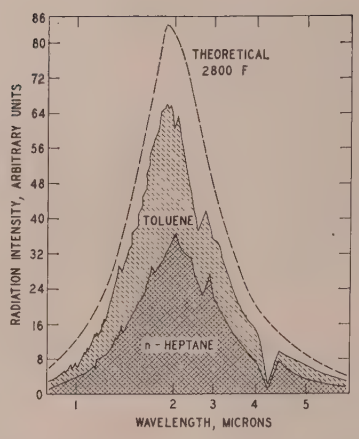


Fig. 2 — Comparison of black body radiant energy with flame emission spectra from 2-in. research combustor at 450 in. of Hg abs

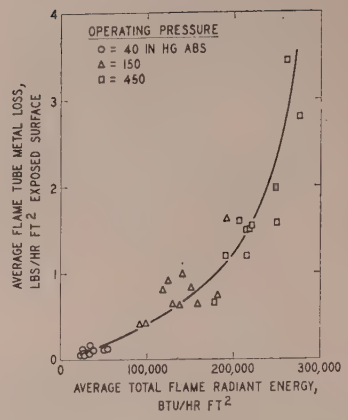

OPERATING PRESSURE
o = 40 IN HG ABS
Δ = 150
□ = 450
AVERAGE FLAME TUBE METAL LOSS, LBS/HR FT 2 EXPOSED SURFACE
AVERAGE TOTAL FLAME RADIANT ENERGY, BTU/HR FT2

Fig. 3 — Effect of flame radiation on durability of 2-in. research combustor

body curve shows that 90% of the total flame radiant energy would be picked up by a suitable pyrometer with sapphire optics, and that less than 1% of the total flame radiant energy is in the visual spectrum. This optical thinness of even luminous flames at atmospheric pressure has made attempts to distinguish between hydrocarbons on the basis of total flame radiant energy very difficult.[14] Even attempts to distinguish between fuels on the basis of variation in combustor metal temperature, when exposed to flames sharply contrasting in visual luminosity, have been unsuccessful at atmospheric pressure.[15]

Recent studies, using a high resolution (Cary Model 14) spectrophotometer, of the visible spectra from isooctane and tetralin when burned in the CRC Luminometer under standard rating conditions have shown only continuous emission spectra, with no evidence of emission bands or of atomic or excited molecular structure. Such a continuum in the visible spectral range of luminous flames at atmospheric pressure is generally attributed to very small particles, small enough to be essentially transparent to radiation of longer (infrared) wavelength.[16,17] Such a spectral distribution of flame radiant energy helps to account for the good correlation found by Randall and Ritcheske[6] between the Texaco lamp, which uses a wide band black body detector (0.3–3 microns), and the otherwise similar CRC Luminometer, with a detector sharply focused in the visible range on the C_2 Swan system (0.48–0.70 microns).

The infrared emission spectra obtained using a laboratory scale combustor operated at high pressure (450 in. of Hg abs) are shown typically in Fig. 2.[12] Such brilliantly luminous flames are characterized by a predominance of continuous, black body type radiation. The average emissivity of this toluene flame approached 1, as shown by its proximity to the theoretical black body curve for this gas temperature. The total radiant energy flux intensity for the toluene flame shown was 160,000 Btu/sq ft/hr, representing about 10% of the total energy released by the combustion process. It is of interest to note that the average emissivity of the paraffinic-type hydrocarbon (normal heptane) flame is quite high, about 0.5, but still only about half of the aromatic-type hydrocarbon (toluene). This emphasizes that the crux of the combustion cleanliness problem under high-pressure operating conditions is control over the size or rate of growth of carbon particles. It also suggests that laboratory studies of this pyrolysis process should be carried out under high pressure.

Combustor Durability

Brilliantly luminous flames have been among the more prominent effects observed in combustion studies at high pressures, that is, above 10 atm. Combustor flame tube deposits (coke) generally drop to negligible levels at high operating pressures in the face of increased carbon formation (smoke).[18] However, reduced depositions frequently are accompanied by buckling of combustor liners, distortion of air inlet holes, and loss of metal, manifestations of high metal temperature.[19]

Investigations of the effect of flame radiant heating on aircraft gas turbine engine durability generally have utilized a measurement of combustor flame tube temperature as an index of durability. Actually, the total heating experienced by the flame tube is a function of both convection from the highly turbulent gas stream and radiation from the flame. A quantitative estimate of the actual fraction of the total heating experienced by the flame tube accounted for by radiation alone in a laboratory scale combustor operating at 15-atm pressure indicated that radiation contributed from 40% with a paraffinic fuel (normal heptane) to 80% with an aromatic fuel (toluene) of the heat flux at the exposed surface of the flame tube.[15] This resulted in a change of about 300 F in flame tube temperature, but its actual significance in terms of combustor durability is unknown. The possible importance of

[14] "Characteristics of Flame Radiation of High Temperature Hydrocarbon Fuels," by W. L. Streets. Summary Report for USAF Contract No. AF 33(616)-5543, Project No. 3048, Task No. 30195. Phillips Petroleum Co. Research Division Report 2307-59R.

[15] "Gas Turbine and Jet Engine Fuels," by G. D. Kittredge. Summary Report for U. S. Navy Contract NOas 58-310-d. Phillips Petroleum Co. Research Division Report 2182-58R.

[16] "Science of Flames and Furnaces," by M. W. Thring. Pub. by John Wiley and Sons, Inc., New York, 1952.

[17] "Spectroscopy of Flames," by A. G. Gaydon. Pub. by John Wiley and Sons, Inc., New York, 1957.

[18] "Preliminary Investigation of Performance of Single Tubular Combustor at Pressures up to 12 Atmospheres," by J. D. Wear and H. F. Butze. NACA RM E53K09, 1954.

[19] "Performance of Tubular Turbojet Combustor at High Pressures and Temperatures," by H. F. Butze and J. D. Wear. NACA RM E55A24, 1955.

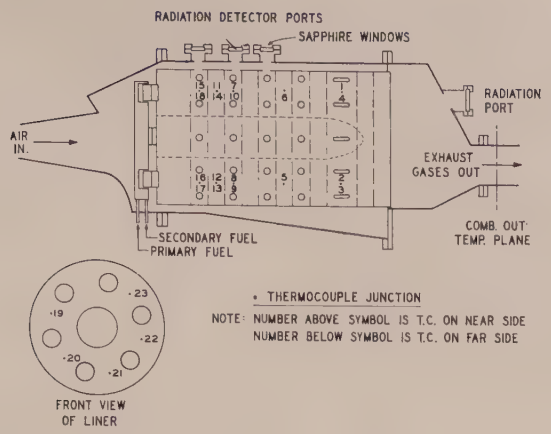


Fig. 4 — Diagram of J-57 single combustor test section

even a few degrees difference in flame tube temperature has been indicated by Hatfield's investigation of high-temperature corrosion of various alloy steel formulations under oxidizing conditions, where sharp transitions or threshold temperatures were found above which the rate of oxidative attack increased sharply, with further increases of a few degrees in metal temperature increasing metal loss rate by an order of magnitude.[20]

It is possible to approach the study of aircraft gas turbine combustor durability using perhaps a more basic measurement, that of flame tube metal burn-off rate, rather than flame tube temperature. The data presented in Fig. 3 are from an investigation of the effect of variations in fuel hydrocarbon structure on flame radiation and combustor durability, conducted in a 2-in. diameter laboratory scale combustor operated under simulated conditions of sea-level take-off, medium-altitude cruising, and high-altitude loitering flight.[14] The flame tubes were fabricated from 2-in. Schedule 40 Type 304 stainless-steel pipe, providing wall thicknesses approaching ⅛ in. Though this is heavier than typical full-scale units, it provided the structural integrity required to sustain the large metal losses experienced during high-pressure operation.

The parabolic shape of the correlation curve between total radiant energy and flame tube metal loss (Fig. 3) serves to emphasize the catastrophic effect of the very high levels of radiant heating encountered at high pressure. The effect of fuel hydrocarbon structure is also clearly evident, as shown by the range in data points at each pressure resulting from fuel variations; for example, an aromatic fuel (toluene) produced a fourfold increase in metal loss over that experienced with a paraffinic fuel (isooctane). However, it was also found that an even greater (order of magnitude) reduction in metal loss could be effected by a change in flame tube metallurgy; in this specific case, by the use of Type 310 stainless steel.

Experimental Program

This paper summarizes a joint investigation by Phillips Petroleum Co., Research Division, and the United States Naval Air Material Center, Aeronautical Engine Laboratory, to determine the effect of fuel chemical composition on the radiant energy from gas turbine combustor flames. This work was sponsored by the United States Navy, Bureau of Aeronautics (now the Bureau of Weapons).

A series of special test fuels was blended by Phillips, rigidly controlling their content as to hydrocarbon type and aromatic structure. Preliminary fuel evaluations were made using the Phillips 2-in. diameter research combustor operated at pressures from 1 to 15 atm for measurement of total flame radiant energy and combustor liner metal temperature, and using the Phillips Microburner at atmospheric pressure for measurement of combustor wall metal temperature and coke and smoke forming tendencies. The performance characteristics of the special test fuels in these laboratory-scale combustors are reported by Kittredge,[15] and will not be discussed in detail in this paper.

Basic data on total flame radiant energy, flame luminosity, combustor liner metal temperature and combustion efficiency were obtained with the total series of 12 test fuels by the Aeronautical Engine Laboratory operating a Pratt & Whitney J57 aircraft gas turbine engine single combustor segment at inlet conditions simulating cruise at altitude for maximum range of a Navy attack bomber. Tests were conducted at three heat input rates of 195, 240, and 285 Btu per pound of air resulting in burner gas temperature rises of approximately 700, 850, and 1000 F. Combustor inlet air pressure was 142 in. of Hg abs, combustor inlet air temperature was 540 F, and combustor airflow rate was 9.6 lb/sec. A Pratt & Whitney J75 aircraft gas turbine engine with afterburner was also operated over its speed range at sea-level static conditions with fuels similar to Test Fuels 7, 8, and 9, to determine their effect on afterburner liner metal temperature.

Valuable contributions were also made to this program by other laboratories. The Naval Research Laboratory, Chemistry Division, determined Luminometer ratings on the total series of 12 test fuels in a room maintained at 71–72 F and approximately 40% relative humidity. The Standard Oil Co. (Indiana), Whiting Research Laboratory, determined Indiana Smoke Point ratings on the total series of 12 test fuels. The Shell Oil Co., Wood River Research Laboratory, evaluated the effect of Test Fuels 1–5, 8, and 9 on the combustor liner metal temperature of a General Electric Co. J79 aircraft gas turbine engine single combustor. Tests were conducted at inlet air conditions of 34 in. of Hg abs pressure, 765 F, and 5150 lb/hr flow rate, with fuel flow rate varied to obtain data at burner temperature rise values from 850 to 1150 F. The latter work in the J79 was sponsored by the United States Air Force, Wright Air Development Center.

Test Apparatus

The J57 combustor liner was instrumented with 23 chromel-alumel thermocouples located at six stations along the length of the liner as shown in Fig. 4. Flame total radiant energy and luminosity were measured at three transverse locations and one axial location, also as shown in Fig. 4. Total flame radiant energy was measured with a Leeds & Northrup Rayotube modified with sapphire optics to cover the infrared spectral range out to 5 microns wavelength as indicated desirable by Miller, et al.[12]

[20] "Heat Resisting Steels," by W. H. Hatfield. *Journal* of Institute of Fuel, Vol. 11, 1958, pp. 254–304.

Flame luminosity was measured by a phototube (RCA 929) with maximum response at 0.4 microns and sensitivity in the visual range of the spectrum (0.3–0.7 microns).

Test Fuels

The entire group of a dozen test fuels involved in this study are described in Table 1. Test Fuels 1–5 were blended specifically to allow evaluation of the relative effects of monocyclic versus polycyclic aromatic fuel constituents on flame radiation produced by JP-5 type jet fuels of the same base ASTM Smoke Point rating. The balance of the test fuels were selected to cover a wide range in Luminometer number to facilitate its comparison with measurements of total flame radiant energy.

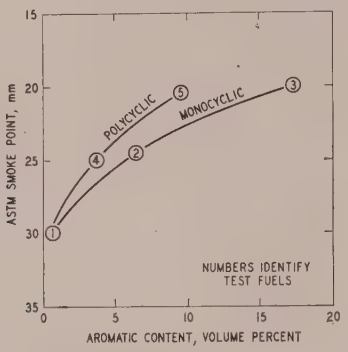


Fig. 5 — Blend characteristics of JP-5 type test fuels

Test Fuel 1 was first prepared by blending a paraffinic (alkylate) fraction with a naphthenic (alkyl decalin — Test Fuel 6) fraction in suitable proportions to provide an aromatic-free, JP-5 type base fuel having a 30 mm ASTM Smoke Point. A monocyclic aromatic (tri-ethyl benzene) and a polycyclic aromatic (methyl naphthalene) were added separately to the aromatic-free base fuel (Test Fuel 1) to obtain monocyclic aromatic Test Fuels 2 and 3 and polycyclic aromatic Test Fuels 4 and 5, having ASTM Smoke Points of approximately 25 and 20 mm. Since both aromatic blending ingredients are of similar (400 F+) boiling range, the volatility of these five test fuels are comparable, differing only a few degrees F in their ASTM Distillation temperatures, as shown in Table 1. The relationship between aromatic content and ASTM Smoke Point for these five test fuels is shown in Fig. 5. Approximately twice as much of the monocyclic aromatic was required to lower the base fuel to either the 25 or 20 mm ASTM Smoke Point level, showing that the lamp is quite sensitive to small concentrations of polycyclic aromatics.

The concentrations of aromatics *as blended* are given in the description of these five test fuels in Table 1, while the concentrations indicated by subsequent inspection tests are listed under hydrocarbon type analysis. It may be noted that the two methods used for determination of both aromatic

Table 1 — Physical and Chemical Properties of Test Fuels

Test Fuel Number	1	2	3	4	5	6	7	8	9	10	11	12
Type	JP-5	JP-5	JP-5	JP-5	JP-5	"HTF"	"JP-150"	JP-4	JP-5	— Commercial Grade —		
Description	Aromatic free base blend	No. 1 plus 6.5% single ring aromatic	No. 1 plus 17.3% single ring aromatic	No. 1 plus 3.8% double ring aromatic	No. 1 plus 9.6% double ring aromatic	Mixed alkyl decalins	Low luminosity kerosine	AEL plant fuel	AEL plant fuel	Isooctane	Normal heptane	Tetralin
ASTM Distillation, F												
Initial Boiling Point	364	367	375	372	375	326	276	136	340	—	200	386
5% Evaporated	381	381	384	382	382	399	282	200	364	211	205	390
10%	386	387	389	386	389	407	282	216	376	—	206	390
20%	392	392	397	392	395	414	284	240	386	—	206	392
30%	396	397	399	397	400	417	284	254	394	—	206	392
40%	400	402	402	402	405	417	284	270	400	—	206	392
50%	405	406	406	407	410	418	284	290	408	—	206	394
60%	411	411	410	412	415	418	284	312	414	—	206	394
70%	417	416	415	419	421	419	284	334	422	—	206	394
80%	426	424	422	428	429	419	286	370	436	—	206	394
90%	442	440	433	444	443	420	286	420	452	—	206	396
95%	470	470	458	472	467	421	288	454	474	—	206	396
End Point	526	522	512	523	516	465	334	492	494	—	252	416
Residue	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.6	0.9	—	0.0	0.5
Loss	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.4	0.6	—	0.0	0.0
Gravity, API	45.9	45.0	43.2	44.3	41.0	34.6	63.1	53.3	41.8	69.5	71.3	14.5
Heat of Combustion, Net, Btu/lb	18,885	18,720	18,574	18,630	18,486	18,384	19,116	18,661	18,563	19,100	19,058	17,243
Viscosity at -30 F, centistokes	12.88	14.68	11.71	14.10	13.36	Solid	—	—	—	—	—	—
Hydrocarbon Type Analysis												
Olefins, Volume % — FIA	3.2	3.2	3.2	3.2	3.2	1.4	1.0	0.4	1.7	—	0.6	1.0
— Bromine No.	1.3	1.3	1.3	1.3	1.3	—	—	—	—	—	—	—
Total Aromatics, Volume % — FIA	2.5	8.6	19.5	4.3	10.2	1.4	1.0	11.7	17.4	—	0.0	94
— Acid Absorption	0.7	4.7	15.9	1.3	6.7	—	—	—	—	—	—	—
Polycyclic Aromatics, Volume % — UV	0.02	0.02	0.03	3.0	8.5	0.01	—	1.2	2.8	—	—	—
ASTM Smoke Point, mm	30	24.5	20	25	20.5	20.5	> 50	27.5	22	39.5	> 50	6
Indiana Smoke Point, mm	44	43	30	43	28.5	27.5	138.5	43	33	80	160	4.5
Luminometer Number	70	59.0	46	57.5	41.5	46.5	136	67	46.5	99.5	180.5	-1.5

and olefin content differ somewhat, both from each other and from the known blending values. The Fluorescent Indicator Absorption (FIA) technique is known to be limited in precision to approximately plus or minus 3 percentage points; therefore, since the ingredients of these five blends are known to be entirely paraffinic, naphthenic, and aromatic, it is believed that indications of olefins by either FIA or Bromine Number are incorrect, as is the indication by FIA of 2.5% aromatics in Test Fuel 1.

Results and Discussion

The detailed data showing the effect of variations in fuel composition on total flame radiant energy, flame luminosity, average and maximum combustor liner metal temperature, and exhaust gas temperature of a J57 combustor are presented in Table 2 for all 12 test fuels with the combustor operating at three different heat input rates. The data are remarkably consistent, showing little or no variation between the 12 test fuels in their relative order of performance regardless of the test condition or the measurement.

The significance of the relationship between total flame radiant energy and flame luminosity in the J57 combustor operating at a pressure of 142 in. of Hg abs was tested by simply plotting the two measurements against each other, which were obtained at the three transverse and the axial position for each of the 12 test fuels at the three operating conditions. This plot is shown as Fig. 6, and the extreme divergence between the two measurements is obvious. This lack of correlation is to be expected from a luminosity measurement representing prob-

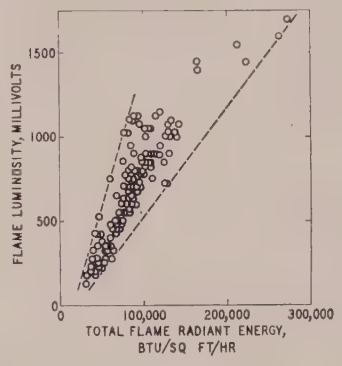


Fig. 6 — Divergence between total flame radiant energy and flame luminosity measurements in J-57 single combustor operating at 142 in. of Hg abs

Table 2 — Effect of Fuel Composition on Flame Radiation and Liner Temperature of J57 Combustor

Test Fuel No.	Total Flame Radiant Energy Btu/ft²/hr × 10⁴				Flame Luminosity, Millivolts				Average Liner Temperature, F							Maximum Liner Temperature	Average Exhaust Gas Temperature
	% Burner Length				% Burner Length				% Burner Length								
	18	33	48	100	18	33	48	100	0	18	25	33	54	80	All		
					ΔT = 700 F			195 Btu/lb of Air									
1	8.9	6.1	3.7	7.4	600	400	200	850	931	742	763	751	685	686	766	1060	1250
2	9.5	7.0	4.0	7.8	725	550	325	1025	931	760	784	774	698	701	782	1070	1245
3	10.8	8.2	4.6	8.4	850	700	425	1100	954	780	806	796	708	705	799	1100	1255
4	9.7	7.1	4.3	8.0	775	550	300	1025	940	765	789	775	702	704	786	1075	1255
5	10.8	7.65	4.6	9.2	900	625	350	1125	990	802	821	810	718	702	816	1125	1255
6	10.7	7.65	4.7	9.2	850	700	350	675	974	785	810	795	705	704	804	1110	1260
7	6.1	4.7	3.3	4.7	350	250	175	525	855	679	697	686	650	679	712	965	1260
8	10.3	7.5	4.0	8.4	850	600	350	1100	950	768	796	784	712	708	793	1090	1255
9	10.8	6.8	4.3	9.0	900	650	300	1125	959	772	813	802	710	709	802	1085	1245
10	7.25	5.6	3.85	6.0	475	350	225	750	889	706	727	719	672	685	738	995	1270
11	5.3	4.3	3.1	4.2	250	200	125	425	839	660	677	669	645	674	697	930	1275
12	22.35	16.45	7.9	12.8	1450	1400	775	725	1175	932	985	991	810	745	954	1360	1255
					ΔT = 850 F			240 Btu/lb of Air									
1	11.0	7.5	5.8	10.2	750	550	350	1000	978	774	800	799	760	737	812	1095	1380
2	10.85	8.2	6.8	10.6	825	750	500	1050	984	799	840	834	792	758	838	1080	1400
3	14.1	9.95	7.4	12.1	1000	800	500	1150	1010	816	856	855	798	759	854	1140	1400
4	12.4	9.25	7.5	10.3	850	700	500	1050	989	789	825	830	780	750	831	1105	1405
5	13.8	10.7	8.7	11.5	1025	850	625	1125	1032	830	872	875	808	767	869	1180	1395
6	12.8	9.8	7.25	11.0	1000	775	450	1050	1000	819	871	849	795	760	850	1140	1385
7	7.1	5.4	4.9	5.05	450	325	225	375	890	704	730	725	712	726	750	995	1405
8	11.6	8.9	7.25	9.6	900	750	550	1075	974	774	823	818	778	748	823	1085	1390
9	13.2	9.8	7.8	10.1	1100	850	550	850	1012	824	855	858	805	768	858	1150	1400
10	8.5	6.7	5.8	6.8	600	450	325	500	929	729	757	758	738	728	775	1035	1425
11	6.3	4.9	4.3	4.5	275	225	175	275	875	681	715	700	692	709	731	975	1405
12	26.3	16.45	8.5	12.8	1600	1450	950	700	1200	972	1014	1030	928	825	1002	1385	1390
					ΔT = 1000 F			285 Btu/lb of Air									
1	10.3	7.8	6.55	6.7	825	600	450	450	968	761	795	800	790	770	815	1120	1519
2	11.75	8.6	7.1	7.4	900	700	550	600	979	776	816	819	812	781	832	1120	1522
3	13.0	10.3	8.2	9.1	1000	900	675	750	1002	804	849	832	838	799	855	1120	1522
4	12.0	9.25	7.5	8.8	950	800	550	750	984	784	824	830	815	782	838	1105	1532
5	14.15	11.0	8.5	9.7	1075	900	650	875	1019	818	863	871	850	802	872	1160	1522
6	13.0	10.2	8.2	8.2	1025	825	550	800	1004	805	845	854	832	792	857	1145	1517
7	6.85	5.7	5.7	4.8	475	350	325	425	901	696	727	728	732	754	757	1015	1545
8	11.6	9.7	8.3	8.5	900	775	625	825	964	786	826	832	828	799	840	1080	1530
9	12.95	11.0	9.5	9.1	1075	900	700	800	1004	806	852	865	855	814	866	1145	1520
10	8.35	6.95	6.85	6.05	575	450	425	475	909	725	759	762	770	774	783	1020	1540
11	5.45	4.7	4.7	3.85	275	250	220	275	885	675	711	712	705	739	740	965	1535
12	27.2	21.25	13.1	8.9	1700	1550	900	400	1199	970	1034	1054	1010	894	1028	1355	1520

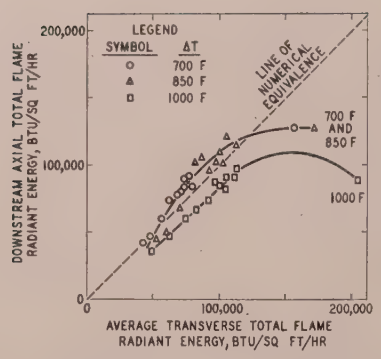


Fig. 7 — Masking effect of combustion products on radiant heating in J-57 single combustor

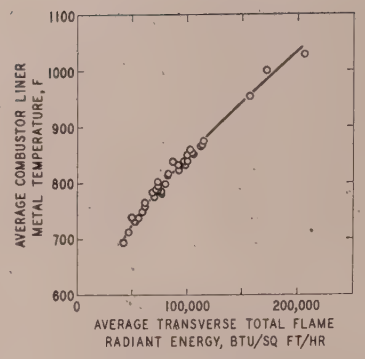


Fig. 8 — Correlation between flame radiation and liner temperature in J-57 single combustor

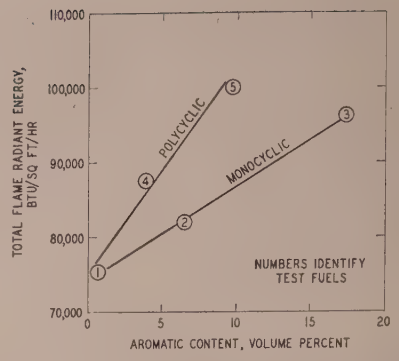


Fig. 9 — Effect of aromatic contents and their structure in JP-5 type test fuels on flame radiation from J-57 single combustor

ably less than 1% of the total radiant energy from the flame, since the large variation in total flame radiant energy observed at essentially constant gas temperatures shows that the flames are not radiating as black bodies with constant emissivity over their entire spectral range. Therefore, the measurements of flame luminosity were dismissed from further consideration, because they are of no significance in the subsequent analysis of the combustor liner temperature data which deals with total heat transfer by flame radiation.

An interesting phenomenon is observed when the average of the transverse total flame radiant energy measurements is plotted against the measurement obtained looking into the combustion chamber along its axis from a downstream location (Fig. 7). The amount of energy radiated by the flame, as measured at the closer transverse positions, increases as the test fuels increase in aromaticity, while the radiant energy received at the downstream location reached a maximum value in intensity. This masking or screening of flame radiation by combustion products has previously been studied in research combustors, where the effect has become so pronounced at operating pressures of 15 atm that paraffin fuels such as normal heptane actually caused more intense radiation from the flame to this downstream location than aromatic fuels such as toluene.[21] The radiant energy reaching either point of measurement reflects the combined action of emission by the high-temperature flame and absorption by the combustion products chilled by incoming air jets. Thus, it appears that the influence of fuel composition on flame radiant heating of metal surfaces surrounding the flame may vary quite markedly depending upon the location of any given piece of metal with respect to the flame. Furthermore, broad generalizations as to the desirability of various hydrocarbon-type structures as fuels, or the undesirability of others as fuel components, with respect to heat transfer by flame radiation, may be quite meaningless in specific engine designs experiencing hot section durability problems.

The essentially straight-line correlation between the averaged values of the transverse total flame radiant energy plotted against average temperature of the J57 combustor liner metal exposed to it for the 12 test fuels at the three different heat input rates is shown in Fig. 8. The correlation between maximum liner temperature and total flame radiant energy while not shown is also good, with maximum temperatures running about 200 F higher than average at low radiation intensities and 400 F higher at high radiation intensities to give a total spread in liner metal temperature of almost 500 F for operation over this range in fuel quality at this condition. The excellence of this correlation prompted the consolidation of the total flame radiant energy values obtained at the three different heat input rates to obtain one overall average value of total flame radiant energy characterizing the performance of each test fuel in the J57 combustor operating at this inlet air condition, which was then available for use in subsequent analysis of the data. While a direct relationship is shown between flame radiation and liner temperature, it is again mentioned that the significance of a given operating metal temperature, and the effect of small changes therein, is unknown in terms of combustor durability until the corrosion and strength characteristics are fixed by the engine designer's specification of the metal for construction.

The effect of variations in the aromatic content and aromatic structure with the five special JP-5 type test fuels on flame radiation from a J57 combustor operating at approximately 5-atm pressure is shown in Fig. 9. It is quite evident, as previously indicated by their ASTM Smoke Points (Fig. 5), that under these operating conditions the polycylic aromatics have a much greater tendency to increase total flame radiant energy than do the monocyclic aromatics. Since the aromatic content of the two series of test fuels were adjusted to give similar ASTM Smoke Points, it is also shown that the J57 combustor is even more sensitive to small concentrations of polycyclic aromatics than indicated by the lamp. These same five test fuels gave no evidence of significant variations in the primary combustion zone liner temperature of a J79 operating at approximately atmospheric pressure, indicating the necessity of high operating pressure to obtain high flame emissivity and significant radiant heating.

The average exhaust gas temperature was essentially constant for all 12 test fuels, indicating no marked variation in combustion efficiency with flame radiation characteristics. There was a slight

[21] "Effects of Flame Radiant Heating on Temperature and Durability of Laboratory Scale Jet Combustor Flame Tubes," by G. D. Kittredge. Special Report for U. S. Navy Contract NOas 58-310-d. Phillips Petroleum Co. Research Division Report 2033-58R.

tendency evidenced for higher combustor temperature rise with Test Fuels 7, 10, and 11, whose flames were lowest in emissivity; however, since these test fuels had narrow and low-temperature distillation curves, this may be a simple expression of engine preference for high fuel volatility. This characteristic of the J57 was probably also evidenced in the J75 by a slight increase in combustion efficiency obtained with "JP-150" as compared with JP-5. More significantly, the afterburner liner dropped about 40 F in maximum temperature and averaged about 20 F cooler with the "JP-150" as compared with the performance of the J75 operating on JP-5. This relatively low level of radiant heating difference between these two fuels, which showed a 135 F average difference in maximum and a 110 F average difference in average combustor liner metal temperature in the J57 which was operated at about 5-atm pressure, is attributed to the relatively low operating pressure in the J75 afterburner, which was between 2 and 2½ atm.

A characteristic relationship between total flame radiant energy from the J57 combustor and ASTM Smoke Point was obtained with the 12 test fuels, as shown in Fig. 10. However, this relationship between flame radiation and lamp smoking tendency exaggerates the significance of high smoke point ratings; a straight-line correlation was obtained between flame radiation and liner temperature. In addition, the ASTM Smoke Point method is limited, as has already been pointed out by others, in its ability to evaluate differences between fuels of low smoking tendency, that is, above 50-mm smoke point. A similar characteristic relationship between flame radiation and Luminometer number is shown in Fig. 11, with the claimed advantage for this method of increased operating range clearly in evidence. The characteristic relationship between flame radiation and Indiana Smoke Point is presented in Fig. 12, showing poorer definition of differences between mid-range fuels and overemphasis of the extremes. Therefore, in the light of the excellent correlation between the present ASTM Smoke Point method and the CRC Luminometer method (Fig. 13), it would seem desirable to consider utilization of the CRC Luminometer as an alternate test method, since it appears to offer certain operating advantages as well as the ability to characterize fuels beyond the limit of the present test method.

In summary, it is emphasized that the visible flame radiation, luminosity, represents little energy at hydrocarbon-air flame temperatures. Carbon particles in hydrocarbon-air flames at atmospheric pressure are not usually of sufficient size to radiate appreciable energy in the significant infrared region. It is with high-pressure hydrocarbon-air flames, where carbon particles can normally grow sufficiently in size to radiate as black bodies in the infrared region, that severe radiant heating is encountered with its accompanying problems of combustor durability. The need to improve aircraft gas turbine engine reliability and performance seems to dictate that this barrier of high intensity flame radiation be moved back; either by more critical specification of fuel hydrocarbon type or by advancements in control over the physical and chemical processes involved in high-pressure pyrolysis of hydrocarbon-air mixtures.

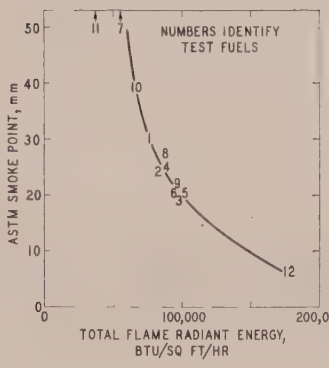


Fig. 10 — Relationship between flame radiation from J-57 single combustor and ASTM smoke point

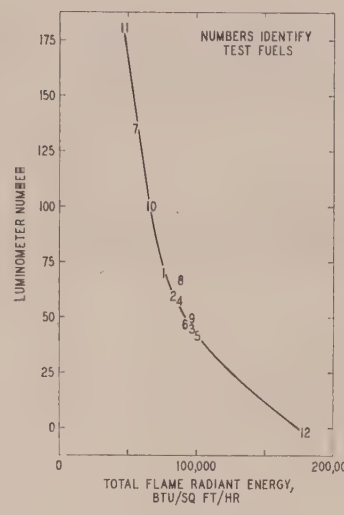


Fig. 11 — Relationship between flame radiation from J-57 single combustor and luminometer number

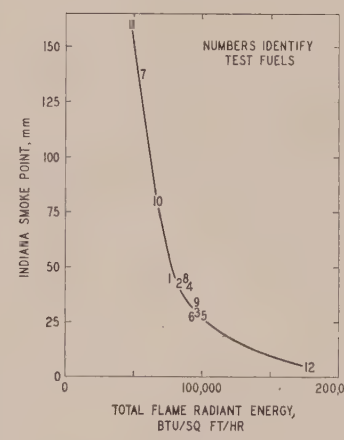


Fig. 12 — Relationship between flame radiation from J-57 single combustor and Indiana smoke point

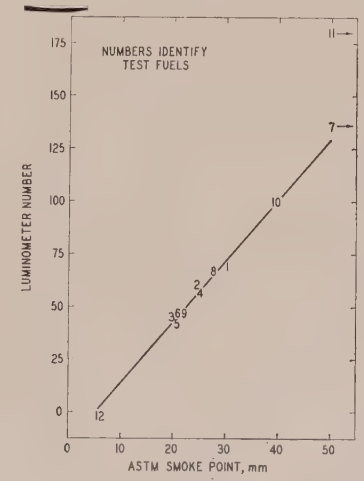


Fig. 13 — Correlation between ASTM smoke point and luminometer number

Physical and Chemical Ignition Delay in an operating diesel engine using

Table 1 — Engine Operating Conditions

Engine Speed, rpm	400 ± 2
Engine Load — Injector rack set to give approximately constant weight of fuel injected per cycle. Engine load varied with fuel	
Inlet Air Temperature, F	82 ± 2
Inlet Air Pressure, in. of Hg	38 ± 0.2
Inlet Jacket Water Temperature, F	160 ± 5
Point of Injection	10 ± ½

Table 2 — Fuel Specifications

Pure Fuels	n-Heptane		n-Cetane	
Specific Gravity, 60 F/60 F	0.6882		0.7773	
Viscosity at 100 F (centistoke)	0.523		3.090	
Boiling Point, F	209.17		548.23	
Cetane No.	57		100	
Commercial Fuels	**Laboratory Fuel**	**40-A**	**U-5**	**T-12**
API, deg	30	35.6	51.0	43.7
Viscosity at 100 F, sec			33.9	32.9
Pour Point, F			−5	B-25
Flash Point, F	150	183	139 (Pensky-Martens)	205 (open cap)
Initial Boiling Point, F	420	400	364	398
10% Point, F	476	426	380	456
50% Point, F	520	512	392	494
90% Point, F	595	603	490	540
End Point, F	620	630	604	586
Cetane No.	40	40	22.5	71.5

IN 1955 three companion research-type papers were presented before the Society of Automotive Engineers.[1,2,3] The first two presented experimental data concerning rates of vaporization, while the third presented a theoretical analysis of the effects of vaporization. This paper presents an extension and continuation of the work discussed by Yu and Wakil. The reader is referred to these references for many of the details of the study.

Apparatus and Experimental Procedure

The apparatus and hot-motored technique used were the same as those described by Yu. Basically, this technique consists of obtaining two successive pressure-time records — the first a normal, fired, cycle; the second, the succeeding cycle but with the injector rack pulled back so that no fuel is injected.

In addition to the hot-motored cycle, a nitrogen cycle was used as described by Yu. Basically, this arrangement scavenges the airbox of the GM 1-71 engine with nitrogen, while the fired cycle pressure-time record is being obtained. Thus, when the inlet ports are opened the cylinder is filled with inert nitrogen for one cycle rather than air. Fuel is in-

* Paper presented at SAE National Diesel Engine Meeting, Chicago, Oct. 29, 1959.

[1] "Physical and Chemical Ignition Delay in Operating Diesel Engine Using Hot-Motored Technique," by T. C. Yu, O. A. Uyehara, and P. S. Myers. SAE Transactions, Vol. 64, 1956, pp. 690–702.

[2] "Combustion Characteristics of Diesel Fuels as Measured in Constant Volume Bomb," by R. W. Hurn and K. J. Hughes. SAE *Quarterly* Transactions, Vol. 6, January, 1952, pp. 24–33.

[3] "Fuel Vaporization and Ignition Lag in Diesel Combustion," by M. M. El Wakil, P. S. Myers, and O. A. Uyehara. SAE Transactions, Vol. 64, 1956, pp. 712–729.

the hot-motored technique—part II

C. W. Chiang, P. S. Myers, and O. E. Uyehara
University of Wisconsin

jected in a normal way during this one cycle. Because of the necessity for complete scavenging of the airbox of all oxygen, it was possible to operate the engine only at low (400) rpm. Considerable time and effort have been spent in development of the technique, but to date the authors have felt that there was neligible oxidation during the nitrogen cycle only at the lower rpm. In this connection it should be pointed out that even if the airbox is completely scavenged, inadequate scavenging of the cylinder with the nitrogen might leave some oxygen in the cylinder.

Experimental Data

The operating conditions of the engine were held constant at the values shown in Table 1, except when specifically treated as a variable. The fuels studied and their specifications are shown in Table 2.

The definitions used are shown in Fig. 1, which is a combination of the three pressure-time records obtained. The definitions are admittedly arbitrary, but are useful for discussion purposes. The end of the injection and vaporization delays were taken as the point at which the curves diverge, as shown in Fig. 1. Experimentally the precision of determina-

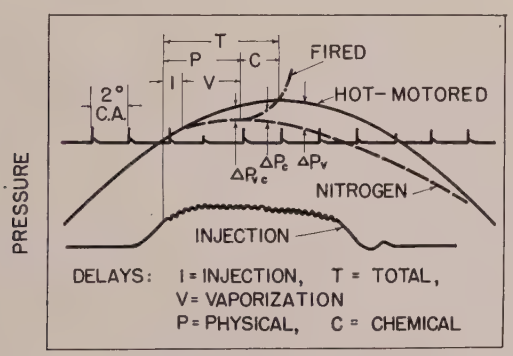


Fig. 1 — Schematic drawing of three pressure-time records

THE PRESENT WORK uses both the hot-motored technique and a nitrogen technique to obtain three pressure-time records — one without either vaporization or chemical reaction, one with vaporization only, and one with both vaporization and chemical reaction.

By comparison of these three records, rates of vaporization and rates of chemical reaction can be determined during the ignition delay period in an operating diesel engine. Such data are shown for different fuels and operating conditions. Estimations are made of the penetration and temperatures existing in the spray.*

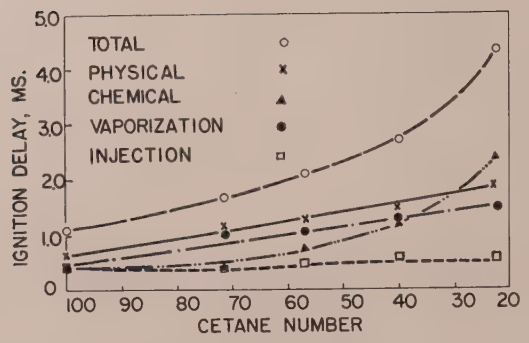


Fig. 2 — Different delays as function of cetane number of fuel

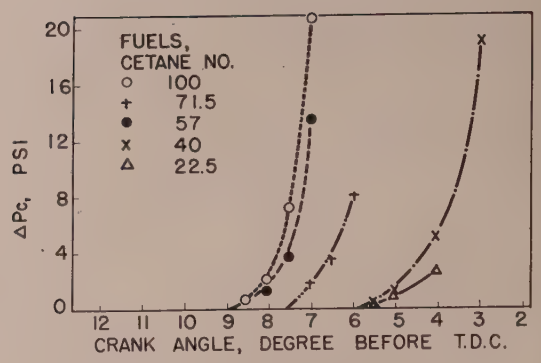


Fig. 3B — ΔP_c versus crank angle for different cetane fuels

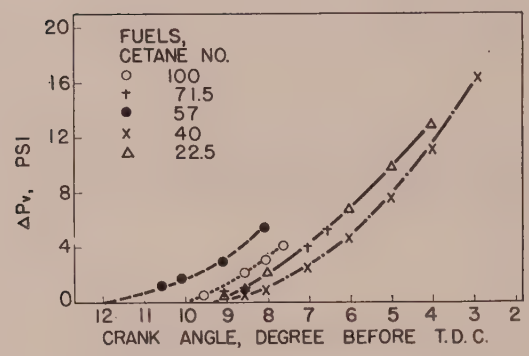


Fig. 3A — ΔP_v versus crank angle for different cetane fuels

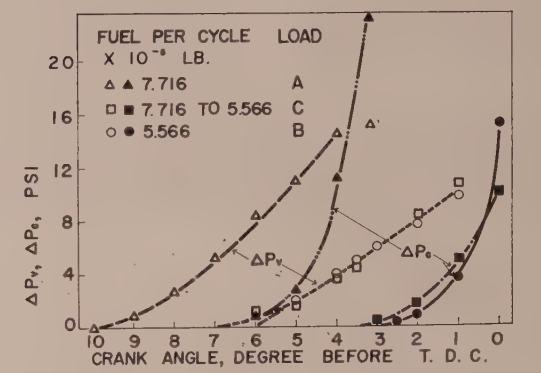


Fig. 4 — ΔP_v and ΔP_c versus crank angle for different engine load and fuel quantity

tion of this point was not too good and could vary at least one-half crank angle degrees between observers. As is shown in Fig. 1, the pressure-time record for the nitrogen cycle eventually approaches the pressure-time record for the hot-motored cycle. This could be caused by differences in heat transfer, by differences in the ratios of specific heat, or by small amounts of chemical reaction. At higher engine rpm, the pressure-time record for the nitrogen cycle would cross the pressure-time record for the motored cycle and have higher values late in the cycle. This was considered as evidence of undesired small amounts of chemical reaction, even though the engine did not fire in the normal sense of the word.

Mention should also be made that it is not clear whether the same density of nitrogen as air at the time of injection or the same pressure of nitrogen as air is desired to give the same spray characteristics. Since it is not possible to obtain both and since experimentally it is easier to maintain the same pressure, this was done.

It will also be noted that Fig. 1 by inference defines the start of injection at the point where small oscillations were noted in the reading of the strain gage on the rocker arm of the injector. Yu[1] defined the beginning of injection as the time at which the strain gage record departed from a straight line. Since the oscillations were believed to be caused by flutter in the check valve, the definition used in this paper is believed more accurate. The use of this definition gives shorter injection delays than those recorded by Yu.

Speaking first of ΔP_v as defined in Fig. 1, if this quantity is to be proportional to the amount of fuel heated and vaporized, the following conditions must be met:

1. There must be very precise cycle-to-cycle reproducibility of engine and instrumentation, because the pressure differences being measured range from 0 to 13 psi out of approximately 500. Even with perfect cycle-to-cycle reproducibility, it is difficult to establish precisely the point at which the nitrogen and fired cycle (Fig. 1) diverge from each other and from the hot-motored cycle. Additional mention is made of this point in the discussion of the accuracy and reproducibility of the results.

2. Heat transfer between the gases and cylinder walls must be the same for both the fired cycle and the nitrogen cycle. Because of the short time interval and the minute differences between the cycles, this condition is believed to be met.

3. ΔP_v is measured when the fuel is sprayed into nitrogen but is assumed to apply when the fuel is sprayed into air. This implies that the spray behaves the same under both conditions. Again because of the minute differences between cycles, this condition is believed to be met.

4. There must be no fuel impinging on the walls of the combustion chamber. When that portion of the cycle under study occurs during the first part of the injection period, this condition is probably met. It may not be met at all times.

5. The partial pressure of the fuel must be negligible. Computations indicate that this would be small

even if all of the fuel were vaporized.[4]

6. There must be no cracking of the fuel. Although computations using normal reaction rates indicate that cracking is negligible,[4] it has been reported a number of times in the literature that oxygen and nitrogen increase cracking rates by orders of magnitude. There is close pressure agreement between the nitrogen and fired cycle during the physical delay period, while the nitrogen content went from approximately 80 to 100%. On this basis cracking was considered to have a negligible effect on ΔP_v.

Since the above requirements were reasonably well met, ΔP_v was used as a measure of the extent of vaporization.

The quantities ΔP_c and ΔP_{vc} must, in general, meet the same requirements — except that the chemical reaction is obviously not zero, since ΔP_c is a measure of the extent of chemical reaction.

Data were first taken for five fuels of different cetane numbers. These were fuels cetane, T-12, n-heptane, 40-cetane lab fuel, and U-5 in Table 2. Fig. 2 presents the data obtained for the different delays plotted against cetane number. Figs. 3A and 3B present plots of ΔP_v and ΔP_c for the same fuels.

It has been noted many times in the literature that for a fixed beginning of injection the ignition delay varied with load at constant speed. This might occur either because engine temperatures change with load or because the quantity of fuel injected changes with load. It has also been noted that in a constant-volume bomb, an increase in quantity of fuel injected increases ignition lag if combustion begins near the end of ignition.[2,3] In an attempt to clarify the relative magnitudes of these effects the engine was operated as if a hot-motored cycle were to be obtained except that the injector rack was not pulled all the way back. In this way engine temperatures corresponded to a heavier load, but the quantity of fuel injected corresponded to a lighter load. Data were also taken at both light and heavy loads using the usual hot-motored technique and laboratory fuel. Fig. 4 presents a plot of ΔP_v and ΔP_c for the three conditions, that is, (1) heavy-load engine temperatures and full charge of fuel, (2) part-load engine temperatures and part-load charge of fuel, and (3) heavy-load engine temperatures and a part-load charge of fuel. Fig. 5 presents ignition delay data with the horizontal scale arbitrarily chosen to give a straight line for the total delay.

In order to simulate the effect of different nozzles, data were next taken by plugging up every other hole in the conventional 6-hole nozzle and also by plugging up three adjacent holes. Fig. 6 presents plots of ΔP_v and ΔP_c for the three arrangments of holes and Fig. 7 presents the ignition delay data. The scale was again arbitrarily chosen to make the total delay plot as a straight line.

In order to establish the effect of inlet air temperature, data were taken at different temperatures using Fuel 40-A. The effect on ΔP_v and ΔP_c are shown in Fig. 8, while Fig. 9 shows a plot of ignition delay versus inlet air temperature.

In order to simulate the effect of supercharging,

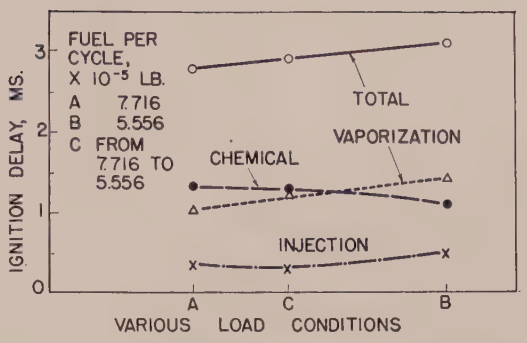


Fig. 5 — Delays for different engine load and fuel quantities

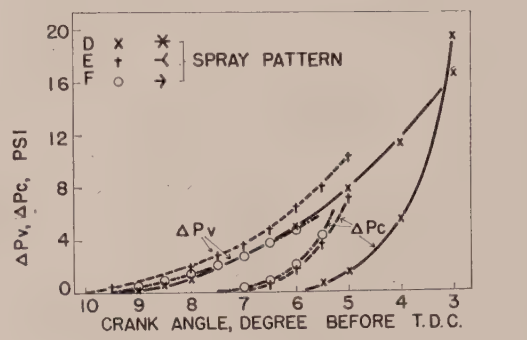


Fig. 6 — ΔP_v and ΔP_c versus crank angle for different nozzle configuration

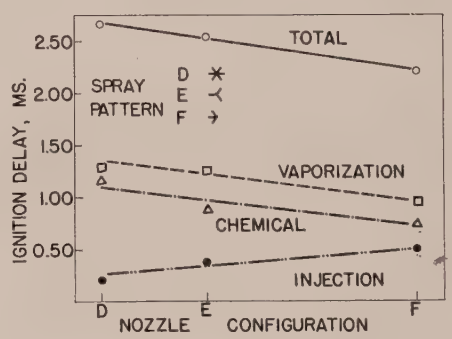


Fig. 7 — Delays for different nozzle configuration

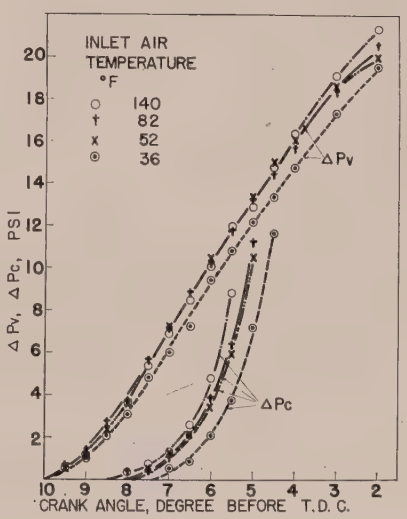


Fig. 8 — ΔP_v and ΔP_c versus crank angle for varying inlet air temperature (Fuel 40-A)

[4] "Pre-Inflammation Studies in Operating Diesel Engine Using Hot-Motored Technique," by T. C. Yu. Doctoral dissertation, University of Wisconsin, 1957.

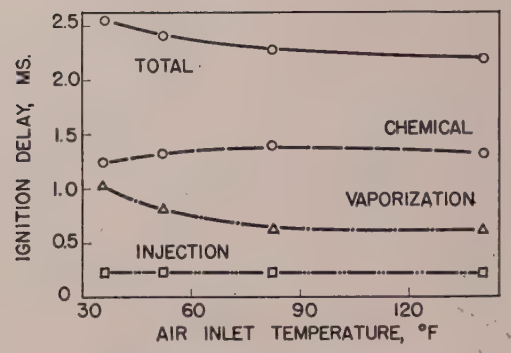


Fig. 9 — Delays for different inlet air temperature (Fuel 40-A)

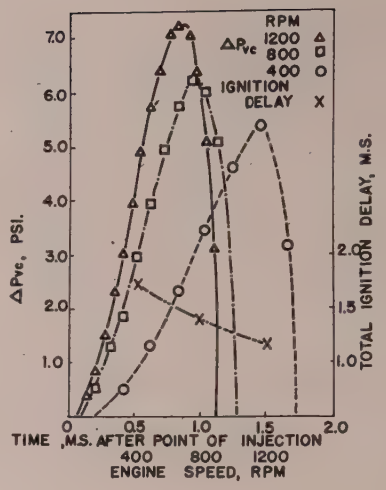


Fig. 11 — ΔP_{vc} versus time and total ignition delay versus engine speed plots

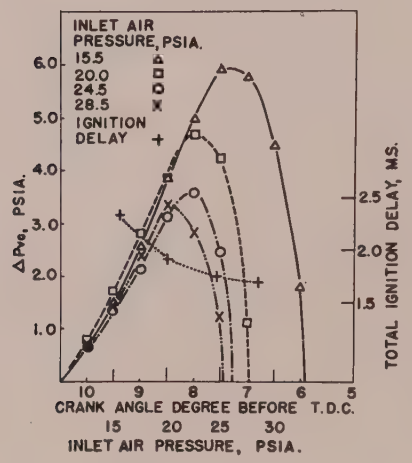


Fig. 10 — ΔP_{vc} versus crank angle and total ignition delay versus inlet air pressure plots

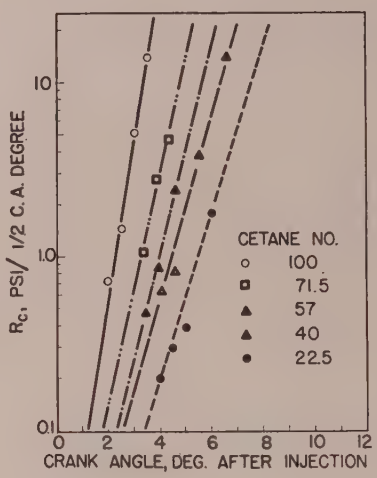


Fig. 12 — Values of R_c versus crank angle for various cetane fuels

data were taken with inlet air pressure as the variable using laboratory fuel. Since the solenoid valves used were inadequate in size to supercharge the engine, the nitrogen technique could not be used at high inlet air pressures; thus, only a plot of ΔP_{vc}, instead of ΔP_v and ΔP_c, could be obtained as shown in Fig. 10. Likewise, only the variation of the total ignition delay could be plotted as shown in Fig. 10.

Similar data were taken with the engine speed as the variable using laboratory fuel. As pointed out before, the nitrogen technique could not be applied to the engine with speeds higher than 400 rpm, because of the necessity for complete scavenging of the airbox. Thus, only a plot of ΔP_{vc} versus time as shown in Fig. 11 could be obtained. The variation of the total ignition delay is also shown in Fig. 11. The speed of the engine was limited to 1200 rpm, because of incapability of the synchronization of the drum camera at higher speeds.

Data were taken by varying the inlet fuel temperature from 50 to 195 F. There was no pronounced effect on the total ignition delay or ΔP_{vc}. The variation of the temperature of the injected fuel leaving the nozzle tip may not be as great as indicated, since the temperature reported is the temperature of the inlet fuel 1½ ft before the nozzle tip.

Discussion of Results

Since the rate of change of the chemical reaction is of interest, values of R_c which was defined as $d\ (\Delta P_c)/dt$ were determined and are plotted in Figs. 12–15. R_c was plotted on a semi-log scale as a function of crank angle degrees after start of injection to eleminate the small variations in point of injection.

The data for the variable cetane fuels show very clearly that not only do high-cetane fuels start to react earlier — their rate of reaction also increases faster than does that of the low-cetane fuel, as evidenced by the steeper slope of the high-cetane fuel in Fig. 12. It must be remembered, of course, that only a very early and minute portion of the combustion process is being studied here and that the situation reverses itself later on when the low-cetane fuel clearly has the higher reaction rate with a consequently higher rate-of-change-of pressure.

Fig. 13 presents similar data for the case where the same part-load quantity of fuel was injected — one time with the cylinder temperatures corresponding to part-load conditions and the other with the cylinder temperatures corresponding to full-load conditions. There is a small difference with the reactions apparently starting a little sooner when the cylinder temperatures correspond to full-load conditions. There is no apparent difference in slopes between the curves. It is interesting to note,

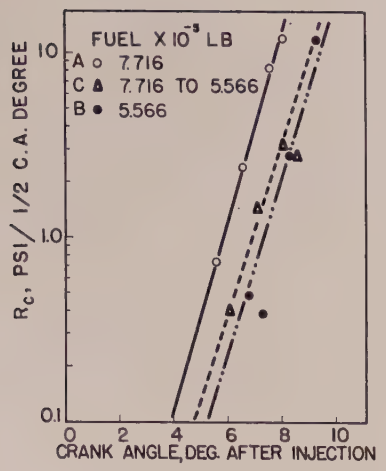


Fig. 13 — Values of R_c versus crank angle for different engine load and fuel quantities

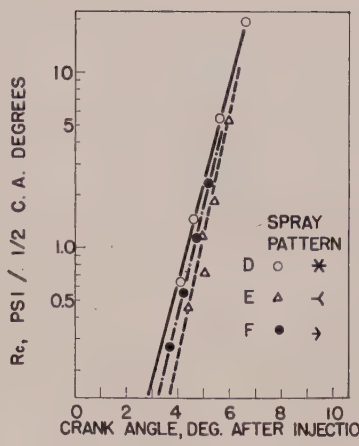


Fig. 15 — Values of R_c versus crank angle for different nozzle configuration

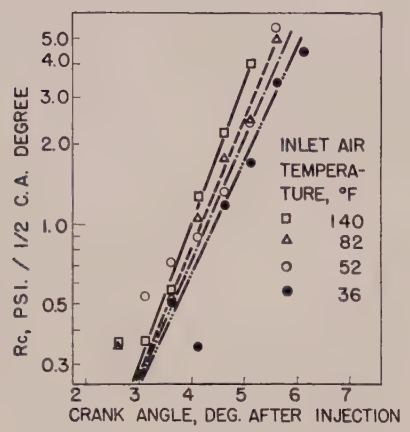


Fig. 14 — Values of R_c versus crank angle for varying inlet air temperature (Fuel 40-A)

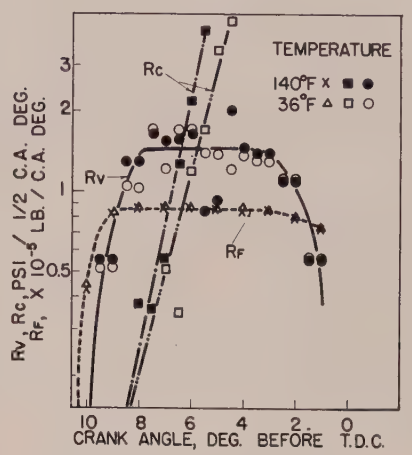


Fig. 16 — Values of R_v, R_c, and R_F versus crank angle for varying inlet air temperature (Fuel 40-A)

however, that the total delay is shorter for the full-load conditions (Fig. 5). This must be primarily due to the increased quantity of fuel injected, since when a smaller quantity of fuel was injected into the same conditions the delay was longer. This is contradictory to results obtained in a constant-volume bomb.[2]

Data for variable inlet air temperatures (Fig. 14) show a consistent trend with the rate of reaction increasing faster with higher inlet air temperatures.

Fig. 15 presents data for the case where some of the holes were plugged. There seems to be a distinct difference in the values of R_c between the three arrangements.

There would seem to be three factors that would cause different values of R_c: (1) an increase in temperature, (2) a change in concentration of the reactants, and (3) a change in the mass of reactants involved. While it is not clear which of these is controlling during this early portion of combustion, the data show that the rate of reaction is increasing approximately logarithmically with time.

Similarly, the rate of change with time of ΔP_v, that is, $d\ (\Delta P_v)/dt$, was defined as R_v. Data for the case where the inlet air temperature was varied were plotted in a similar manner (Fig. 16). For comparison purposes the rate of fuel injection R_f and R_c are also shown. While the curves for R_v are not smooth, it does appear that soon after the rate of fuel injection R_f, (as indicated by a strain-gage on the rocker arm) became constant R_v also became constant. This could indicate complete and rapid vaporization of the fuel but since high-speed photographs[5] show the presence of liquid particles all during the delay period, it is thought more likely to indicate adiabatic equilibrium as suggested by Wakil.[3]

Attempts were made to plot corresponding values of R_v for the other conditions but considerably more spread in the data were obtained. Values of ΔP_v can be obtained either of two ways: Method 1 where it is measured directly as indicated in Fig. 1; and Method 2 where ΔP_{vc} and ΔP_c are each measured and ΔP_v obtained by difference. A critical analysis of the data showed that, while both are correct theoretically, from a practical standpoint Method 1 was the most accurate. Unfortunately, most of the data were not taken so that ΔP_v could be obtained by Method 1.

Estimations were previously made of the amount of heat required to vaporize the fuel completely and superheat the vapor to the initial temperature of the air.[1,2] Both sets of data (as well as similar estimations made from the present data) show that

[5] "Slow-Motion Study of Injection and Combustion of Fuel in Diesel Engine," by C. D. Miller. SAE Transactions, Vol. 53, 1945, pp. 719–733.

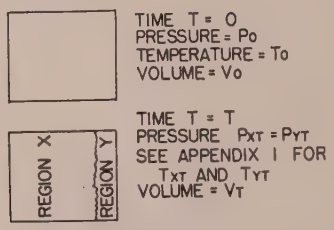


Fig. 17 — Schematic diagram for thermodynamic analysis

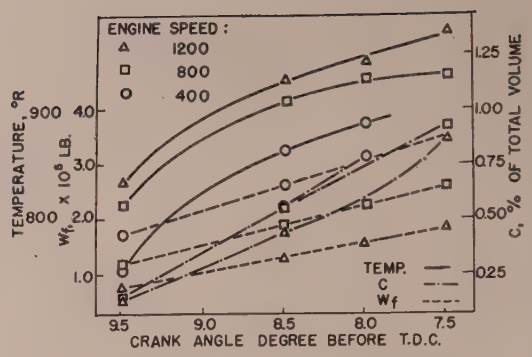


Fig. 20 — Temperature, W_f, and per cent volume versus crank angle for different speeds

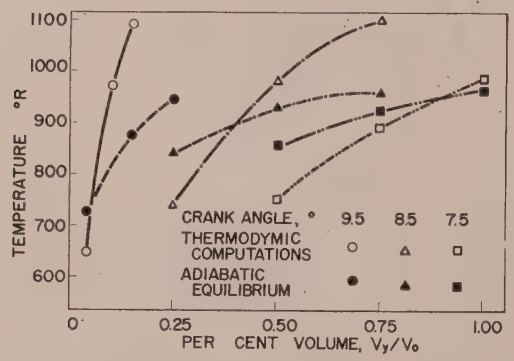


Fig. 18 — Temperature versus per cent volume for cetane

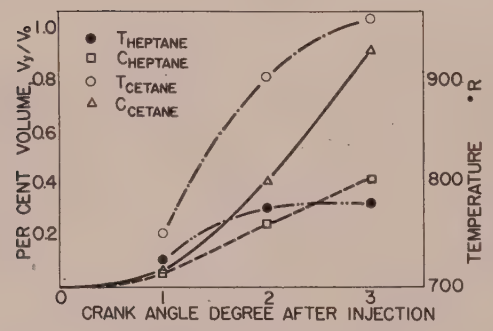


Fig. 19 — Temperature and per cent volume versus crank angle for cetane and heptane

only 60–75% of this heat comes from the air. This could either mean that the additional heat required was either supplied from other sources (impingement) or not supplied at all. In an attempt to determine which of the above viewpoints is correct, a thermodynamic analysis of the physical arrangement shown in Fig. 17 was made. This arrangement simulates the situation during early injection where a small portion of the air in the combustion chamber is cooled by the spray, causing a decrease in the total pressure while the main portion of the air is undisturbed by the spray. For simplicity the temperature of the undisturbed air and the volume of the container were assumed to remain constant. Since only a few crank angle degrees near top dead center are involved, this is a reasonable approximation. The Appendix gives the relationships used in the analysis. Basically, the equations say that because Region y is cooled the pressure drops with the exact value being dependent on the temperature to which Region y is cooled as well as the ratio V_y/V_o, that is, the relative size of V_y and the initial conditions.

Using the experimentally determined values of ΔP_v for the value of $P_o - P_{xt}$, that is the drop in pressure due to cooling Region y, a curve can be plotted of V_y/V_o versus T_{yt}, the temperature of the cooled region for the specified initial conditions. Such a curve is shown in Fig. 18. If a particular value of V_y/V_o is chosen, the weight of air in V_y can be determined. Since the weight of fuel injected up to this point is known if the fuel is assumed to be concentrated in the cooled region, the air fuel ratio is then known. Using this air fuel ratio and adiabatic equilibrium computations (Fig. 12, Wakil[5]), the temperature that would exist in Region y if adiabatic equilibrium existed can be determined and plotted in the same cooordinates (Fig. 18). At the crossing point of these two curves both the condition of adiabatic equilibrium and the experimentally determined value of ΔP_v ($P_o - P_{xt}$) are satisfied.

Repeated computations for the pure fuels cetane and heptane (assuming an initial fuel temperature of 200 F) produced the curves shown in Fig. 19 where the ratio V_y/V_o and the temperature at the crossing point of the curves (shown schematically in Fig. 18) are plotted as a function of crank angle. Since the computations include only the cooling effect of the fuel they were deliberately restricted to the region where $\Delta P_{vc} = \Delta P_v$ (Fig. 1), that is, $\Delta P_c = 0$.

Fig. 19 very clearly shows that the temperature of the cetane spray is higher than the temperature of the heptane spray. While these different temperatures are solely a result of the physical characteristics of the fuels, the speed of the reactions that occur at these temperatures are the result of the chemical characteristics of the fuel. This emphasizes again the interrelationship in the diesel engine between the physical and chemical characteristics of the fuel.

Fig. 20 shows that the temperature of the spray for higher engine speed is higher than that for lower engine speed. This is because there is less fuel injected per crank angle degree and there is less time for heat transfer at the higher engine

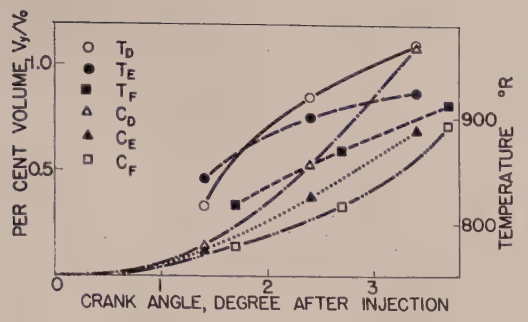


Fig. 21 — Temperature and per cent volume versus crank angle for different nozzle configuration

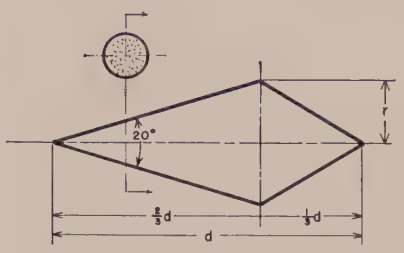


Fig. 22 — Schematic drawing of spray shape

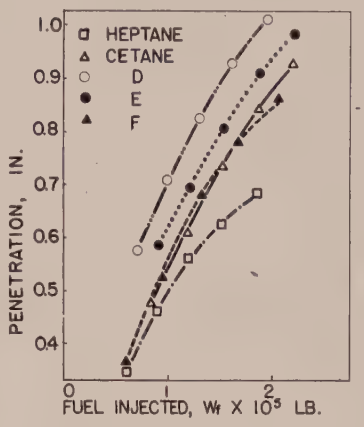


Fig. 23 — Computed spray penetration versus crank angle for different fuels and nozzle configurations

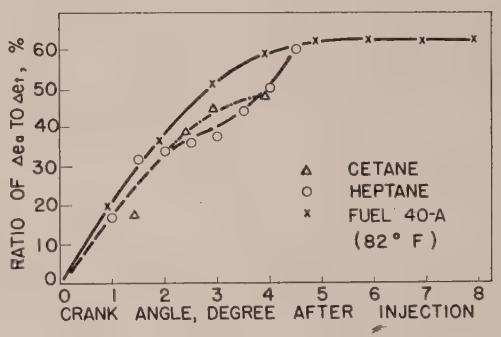


Fig. 24 — Ratio of Δe_a to Δe_t versus crank angle for three different fuels

speed. The volume ratio c at 400 rpm shows a higher value than that at 1200 rpm because of a larger quantity of fuel injected into the combustion chamber at a given crank angle, as shown in Fig. 20. From the plot of the volume ratio c, it is obvious that the penetration for the low engine speed will give a higher value than that for higher speed.

Although the data of Fig. 12 (Wakil[3]) are really not applicable to a complex fuel, similar computations were made for the data when some of the holes in the nozzle were deliberately plugged, on the basis that the results should be approximately correct relative to each other. The results are shown in Fig. 21. The computations show that with no holes plugged higher temperatures are obtained than with half the holes plugged. The data of Fig. 15 show that chemical reaction starts the earliest in this case as would be expected. However, Fig. 15 also indicates that in the case of the unplugged holes the rate of chemical reaction does not increase as rapidly for the unplugged holes in spite of its earlier start. The reasons for this are not clear.

Durant and Eltinge (Fig. 7)[6] shows data on smoke when three holes were plugged. The data at 60% load would probably be most nearly comparable to the present data It is interesting to note that with three holes plugged the fuel is concentrated into a smaller volume, that is, V_{yt}/V_o is smaller at the same crank angle and more smoke is obtained.

From various high-speed photographs the shape of the spray coming from a hole was estimated (Fig. 22). Assuming that the spray from each hole formed a cone having this shape with the apex at the hole, the penetration of the spray giving the required ratio of V_y/V_o can be determined.

Fig. 23 presents computed values of spray penetration using the method outlined above. The values so determined seem quite reasonable both with respect to magnitude and relative values. Heptane is a nonviscous fuel and undoubtedly has less penetration than the more viscous fuel cetane. When three holes are plugged, the injection pressure must increase since the same quantity of fuel must flow through a smaller area in the same length of time. This increased pressure will give better atomization, which will tend to cause less increase in penetration than might be expected due to higher injection pressure. Both 3-hole arrangements gave somewhat higher penetrations, although the reasons for the differences between the two 3-hole arrangements is not clear.

Estimations were next made of the quantity of heat required to change 1 lb of the fuel from a liquid at 200 F to a vapor at 1000 F (the initial temperature used for the thermodynamic computations). This value was called Δe_t. Computations were next performed on the quantity of heat required to change 1 lb of fuel from a liquid at 200 F to the adiabatic equilibrium temperature (Figs. 19 and 20). This value was called Δe_a. Fig. 24 presents the results of these computations for three different fuels.

Fig. 24 can be compared with Fig. 16 of Yu[1] if the vertical scale of Fig. 16 were plotted as a percentage. Fig. 25 presents similar data taken by Selden and Spencer[7] using a constant volume bimb with nitrogen. The agreement in numerical magnitude between the bomb and engine is excellent. The small variation of the ratio with pressure is considered

[6] "Fuels, Engine Conditions, and Diesel Smoke," by J. B. Durant and Lamont Eltinge. Paper presented at SAE Annual Meeting, Detroit, January, 1959.

[7] "Heat Transfer to Fuel Sprays Injected into Heated Gases," by Robert F. Selden and Robert C. Spencer. NACA Report No. 580, 1937, pp. 91–107.

significant since Lee shows that the penetration at any given time decreases with bomb pressure.[8] Thus, if impingement cause the ratio to be less than 100%, the ratio should increase appreciably with an increase in bomb pressure. It should be noted, however, that Selden and Spencer indicate the belief that impingement was the cause of the low ratio. The agreement between the computed data and the experimental facts is excellent, and it seems to the authors that it must be accepted that adiabatic equilibrium exists in the main part of the spray.

If this condition is accepted as existing it follows then that adiabatic saturation and not impingement is the cause for the low (60–75%) ratio of heat-supplied-to-fuel to heat-required-to-vaporize the fuel. This conclusion is also strengthened by the observation that in the region under consideration a maximum of approximately one-half the fuel required for 60–75% of full load has been injected. If the impingement is appreciable at this point, it would mean that impingement is much greater than is normally considered present.

Conclusions

1. All evidence points to the presence of adiabatic equilibrium in the spray soon after injection starts. Adiabatic equilibrium must cease to exist, however, in the edges of the spray.
2. Different fuels and different nozzle configurations give different spray characteristics. These different spray characteristics result in different local air fuel ratios and, thus, affect the adiabatic equilibrium temperature.
3. Since the physical characteristics of the fuel affect the adiabatic equilibrium temperature and the spray characteristics and since chemical reaction rates are markedly affected by temperature, it follows that the physical characteristics of the fuel affect the chemical delay as well as the physical delay.
4. After chemical reaction has started there are differences between fuels in the rate at which the chemical reaction increases. This rate of increase is higher for high-cetane fuels and it is undoubtedly the cause of the less harsh combustion experienced with the high-cetane fuels. On the other hand, low-cetane fuels seem to have comparatively slower early reaction rates followed by very rapid reactions and harsh combustion.
5. All evidence indicates that for the small period studied (less than one-half the injection period) impingement is negligible and the fact that less heat is extracted from the air than is theoretically required for complete vaporization and superheat is due to the existence of adiabatic equilibrium.

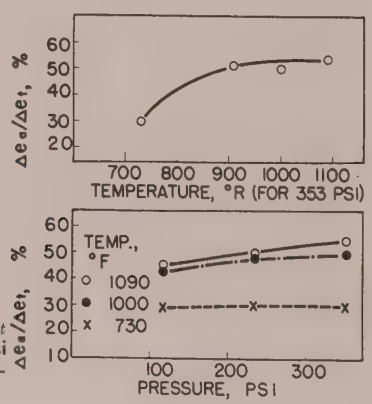


Fig. 25 — Ratio of Δe_a to Δe_t versus temperature (for 353 psi pressure) and different pressures

Acknowledgments

The authors wish to acknowledge the support of the following organizations.

Equipment from the Research Laboratories, General Motors Corp. Fellowship support equipment and salary support from the College of Engineering, University of Wisconsin, and the Wisconsin Alumni Research Foundation. Dr. M. M. El Wakil of the University of Wisconsin contributed to analysis of data.

APPENDIX

Assuming a perfect gas and by definition of the physical situation:

$$P_o V_o = W_o R T_o \tag{1}$$

Since Region y is cooled it follows that Region x expands and that the pressure drops. For Region x:

$$\frac{T_o}{T_{xt}} = \left(\frac{P_o}{P_{xt}}\right)^{\frac{k-1}{k}} \tag{2}$$

Applying the perfect gas relationship to Region x and Region y, using Eq. 2, and remembering $P_{xt} = P_{yt}$:

$$\frac{V_{xt}}{V_{yt}} \quad \frac{W_{xt}}{W_{yt}} \frac{T_o}{T_{yt}} \left(\frac{P_{xt}}{P_o}\right)^{\frac{k-1}{k}} \tag{3}$$

also:

$$\frac{V_{xt}}{V_{yt}} = \frac{V_o}{V_{yt}} - 1 \tag{4}$$

and

$$\frac{W_{xt}}{W_{yt}} = \frac{W_o}{W_{yt}} - 1 \tag{5}$$

Combining Eqs. 3–5 gives:

$$\frac{V_o}{V_{yt}} - 1 = \left(\frac{W_o}{W_{yt}} - 1\right) \frac{T_o}{T_{yt}} \left(\frac{P_{xt}}{P_o}\right)^{\frac{k-1}{k}}$$

$$= \left(\frac{P_o}{P_{xt}} \frac{T_{yt}}{T_o} \frac{V_o}{V_{yt}} - 1\right) \frac{T_o}{T_y} \left(\frac{P_{xt}}{P_o}\right)^{\frac{k-1}{k}} \tag{6}$$

Let :

$$C = V_{yt}/V_o$$

$$A = T_o \left(\frac{P_{xt}}{P_o}\right)^{\frac{k-1}{k}}$$

$$B = \frac{P_o}{T_o P_{xt}}$$

If the initial conditions are specified and if the value of ΔP_v is known it follows that A and B are known constants. Substituting these definitions into Eq. 6 gives:

$$\frac{1}{C} - 1 = \left(\frac{B\, T_{yt}}{C} - 1\right) \frac{A}{T_{yt}}$$

Rearranging gives:

$$T_{yt} = \frac{A\, C}{A\, B + C - 1}$$

Thus, a curve of T_{yt} versus C may be plotted.

[8] "Comparison of Fuel Sprays from Several Types of Injection Nozzles," by Dana W. Lee. NACA Report No. 520, 1935, pp. 251–288.

REFRACTORY NONMETALS for use in hypersonic aircraft

John M. Nowak
Bell Aircraft Corp.

THIS PAPER describes screening tests by Bell Aircraft on ceramic and graphite materials for hot-spot areas. Results show that the best oxidation resistance at 2800 F was exhibited by a refractory body consisting of 90% silicon carbide and 10% boron carbide. Graphite coated with silicon carbide shows promise for applications at temperatures of at least 2800 F, and possibly 3700 F, for short periods.

The purpose of the tests was to determine the merits of:

1. Ceramic and graphite materials for hot-spot areas.
2. Refractory inorganic materials as thermal insulants.

The refractories were being considered for the wing leading edges and the nose of a proposed hypersonic aircraft. The insulants were for application between the outer skins and the interior structure.*

UNTIL VERY recently, with the relatively slower vehicles, metals have adequately fulfilled the structural requirements in aircraft and missiles. However, as the speeds in our new aircraft and missiles continue to Mach 10, Mach 15, and beyond, use of conventional designs radically limits the use of present metals. Metal alloys useful to temperatures of about 1350–1600 F are commercially available. Some of the high-temperature alloys, as an example a high cobalt type (Haynes Alloy 25), have reported stress rupture strengths in wrought form of 7200 psi at 1800 F after 100-hr exposure at 1800 F. The short time tensile strength of this alloy is reported to be about 22,700 psi at 1800 F. However, as operating temperatures increase to 1800 F and above, more and more of our designers are looking to ceramics and other nonmetallic refractory materials to solve their material problems.

The use of ceramics in structural and semistructural applications in high-speed aircraft and missiles is being developed as a natural alternate for the conventional and refractory metals because of the temperature limitations and/or lack of oxidation resistance inherent in the latter. Ceramics is one of the oldest of arts. Yet it is only recently that modern technology has been extensively applied to the study of ceramics. Conventional ceramic materials are prepared by mixing certain amounts of plastic, filler, and flux. The ceramic material processed and prepared for the forming operation had ingredients which function in these three ways. It had to be plastic so that it could be formed into any

* Paper presented at SAE National Aeronautic Meeting, New York, April 2, 1959.

Table 1 — Mechanical Properties of Aluminum Oxide and Molybdenum
Ultimate Tensile Strength (psi)/Density (lb/in.^{3})

Temperature F	Aluminum Oxide (Al_2O_3)[a]	Molybdenum (0.5% Ti-Mo)[b]
70	264,000	—
570	250,000	—
1470	236,000	250,000
1920	229,000	200,000
2070	215,000	170,000

[a] Tensile strength data from "Electrical & Technical Ceramic Industry of Germany PB-18776," Mapleton House Publishers, Brooklyn, N. Y., August, 1947. The theoretical density of 0.144 lb/in.^{3} was used in calculation of the ratio.

[b] Data from "Molybdenum for High Strength at High Temperatures," by Robert R. Freeman and J. Z. Briggs. Climax Molybdenum Co., reprinted from *Jet Propulsion*, Vol. 27, Part 1, February, 1957.

shape; it had to have a low drying shrinkage so that cracks would not develop which result in failure of the ware; and it had to have low enough fusibility to be fired into a strong product at temperatures economically practical. The minerals which contribute the property of plasticity are called plastics in the ceramic field of technology. They are not the same materials to which the term plastic has now been generally associated. The minerals which contribute the property of fusibility are called fluxes. Minerals which contribute the property of shrinkage control are called fillers. However, many of *today's* ceramic materials, such as pure aluminum oxide, are not compounded in the conventional manner. High-purity bodies may be formed by hydrostatic pressures and high temperatures. Bodies may be fired at very high temperatures. Some recently developed commercial bodies, such as dense silicon carbide, do not use a flux but are self-bonded. Ceramic materials have been evaluated by our laboratories primarily for application in the hot-spot areas of advanced hypersonic aircraft.

Brittleness is the major drawback in using ceramic materials as structural hardware. Although the materials are usually poor in tensile strength at room temperature, the compressive strength is generally excellent. Many ceramic materials, however, have better tensile strengths than metals at temperatures of 1800 F and above. James Gangler and associates[1] reported a short time tensile strength of about 22,500 psi at 1800 F for a boron carbide body. Designs using a ceramic material should take into account the relatively high tensile strengths of some of these materials at elevated temperatures. Studies are now underway by Dr. Joseph Pask and associates at the University of California, and probably others, to improve the ductility of ceramics. It may be many years, however, before ductile ceramics are developed. In the meantime, studies are underway in our laboratories and elsewhere to utilize in flight vehicles the better heat-resistant properties of ceramics. Some ceramic materials have what appears to be definite strength advantages over metals at elevated temperatures. We have compiled Table 1 showing the strength-to-weight ratio of an alumina body compared to molybdenum at various elevted temperatures. It can be seen that alumina can be considered competitive in the temperature range of about 2000 F, based on ultimate strength-density ratio. Metals for use at temperatures of 2000 F and higher, possessing high mechanical strength, are subject to oxidation. Reviewing this data, there are, of course, many other factors which must be considered, including: effect of brittleness of the alumina on design application, the creep rupture strength, the effect of oxidation on molybdenum, effect of coatings on the strength of molybdenum, and the like. This data is presented to bring to your attention the possibilities that do exist for ceramics as structural materials in hypersonic air vehicles even at the relatively low temperatures of 2000 F.

The use of refractory inorganic materials as high-temperature, low-density insulants has been accepted by industry for many years. Thus, quartz fibers and alumina-silica fibers have found extensive applications in commercial and military areas. Although a number of properties of these materials have been reported from time to time, the effect of vibration on these materials at high temperatures has not been investigated to date to our knowledge. Vibration at elevated temperatures is a condition which will undoubtedly be experienced by the high-speed advanced aircraft now being designed. We hope that our presentation will add to the sum of knowledge on the possible application of ceramics and other refractory inorganic materials to high-speed flight vehicles. One cannot overlook any of the materials available. Our studies were directed toward the use of ceramics and the refractory materials in applications at elevated temperatures for periods of 1–100 hr at temperatures of 1800 F and above.

Results

About six years ago our laboratories undertook the investigation of materials for application to proposed hypersonic aircraft. Reviewing the material problems which we could anticipate at that time and based on design information available, we decided to direct our preliminary laboratory efforts into several functional channels. Our studies were divided into the following: (1) ceramic and graphite materials for hot-spot areas, (2) refractory inorganic materials as thermal insulants, (3) high-temperature metallic alloys, and (4) coatings for high-temperature metallic alloys and nonmetals. In the proposed vehicle, the refractories were considered for the wing leading edges and the nose. The insulants were for application between the outer skins or walls and the interior structure. The refractory metals were being considered as outer walls in the lower temperature areas of the vehicle. The coatings were to impart oxidation protection to the metals and nonmetals at elevated temperatures. The results of our coating evaluation studies and the studies on the refractory metals are not included in this paper.

Refractory Materials for Hot-Spot Areas — Tests were conducted on a number of selected high-temperature ceramic and graphite materials which may be useful as refractories for hot-spot areas where temperatures in excess of 2500 F and approaching 4000 F may be anticipated. Most of the high melting point materials have mechanical, thermal, or chemical properties which make them unsuitable for some of these applications. Although

[1] "Physical Properties at Elevated Temperatures of Seven Hot Pressed Ceramics," by J. J. Gangler, C. F. Robards, and J. E. McNutt. NACA TN 1911, July, 1949.

no material to our knowledge is presently available which will meet the target requirements for all hot-spot areas, test data establishing useful life temperatures for available materials is imperative. A number of refractory carbide and graphite materials were selected on the basis of their theoretical properties and were screened for rate of oxidation at 2800 F under a controlled oxygen flow of 1 liter per min. This temperature was selected as a plateau representing the lower temperature limits of proposed applications.

Oxides as a class appear to have lower useful operating temperatures than do the carbides. The oxides tend to have lower strength at the high temperatures. Some are subject to thermal shock. The carbides, however, generally have higher melting points, higher strengths at the elevated temperatures and are not subject to the degree of thermal shock that is common to some oxides. At sufficiently elevated temperatures the carbides will oxidize. One of the purposes of our studies on the carbides and graphite was to ascertain the degree of oxidation which we could expect.

In addition, preliminary temperature tests to 3700 F were conducted on a magnesia-zirconia body, on a dense silicon carbide body, and on a zirconium carbide body. Modulus of rupture tests were conducted at temperatures to 2700 F on an experimental, aluminum modified, dense silicon carbide body, on a commercially available dense silicon carbide body, on a hot pressed 90/10 silicon carbide-boron carbide body, on a slip cast 90/10 silicon carbide-boron carbide body, and at temperatures to 1500 F on a number of zirconia bodies. The aluminum modified silicon carbide was fabricated by a refractory manufacturer in an attempt to obtain a commercially producible silicon carbide possessing, a modulus of rupture approaching the best hot-pressed variety known to us. A hot-pressed, high-strength SiC body with a reported modulus of rupture of 70,000 psi at 2500 was prepared by Alfred University on a laboratory scale.

Of the bodies tested only SiC-B_4C, recrystallized SiC, and AGR graphite coated with about 0.005 in. SiC were capable of withstanding exposure to oxygen at 2800 F for 3 hr without excessive deterioration. The ZrC oxidized to a powder. The TaC and the TaC-ZrC bodies oxidized excessively. Based on the change in weight and the amount of CO_2 collected per square centimeter of surface area, the SiC-B_4C body showed the greatest resistance to oxidation under those test conditions. Examination of the SiC coated graphite showed pronounced voids under the coating. These voids were concentrated at the edges. This would indicate that it may be difficult in getting an adequate coating on the edges. It may also indicate that the coating tends to crack at the edges due to difference in thermal expansion of graphite in different planes. Detailed oxidation resistance test data for the bodies evaluated is given in Table 2. The weight of CO_2 collected per square centimeter of exposed area against exposure time for all samples tested in a controlled oxygen flow is plotted in Figs 1–3.

To determine resistance to higher temperatures, preliminary qualitative tests consisting of exposing

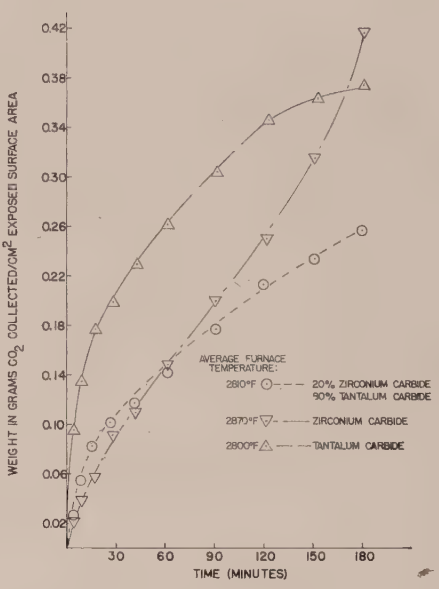


Fig. 1 — Weight CO_2 (grams/cm^2) surface area versus time

Table 2 — Oxidation Resistance of Various Refractory Carbide and Graphite Bodies[a]

Material	ZrC	TaC	SiC-B_4C	ZrC-TaC	Recrystallized SiC	AGR Graphite Coated with SiC
Surface Area, cm^2	18.01	9.21	15.80	4.23	33.09	21.19
Volume, cm^3	4.21	1.57	3.00	0.57	6.43	7.42
Weight, grams	24.0208	17.9954	9.1089	5.5541	16.2308	13.3738
Apparent Density, grams/cm^3	5.70	11.42	3.02	9.67	2.52	1.80
Apparent Porosity, %	6.24	14.59	0.35	18.12	20.71	8.54
Weight after Acid Wash, grams	24.0191	17.9950	9.1075	5.5529	16 2141	13 3697
Loss of Weight from Wash cycle, grams	0.0017	0.0004	0.0014	0.0012	0.0167	0.0041
Weight before Heat-Treatment, grams	24.0191	17.9950	9.1075	5.5529	16.2141	13.3697
Weight after Heat-Treatment, grams	Yellow powder	Gray fused mass	9.1149	Buff colored mass fused to boat	16.5733	13.1146
Gain in Weight after 3-hr exposure, grams	—	—	0.0074	—	0.3592	0.2551 (loss)
Average Temperature of Furnace, F	2870	2800	2840	2810	2820	2810
Weight of CO_2 Produced after 3-hr exposure, grams	7.4614	3.4499	0.0259	1.1110	0.1697	0.9865
Theoretical Weight of CO_2 Producible	10 2400	4.1057	7.2493(B_4C) 10.0030(SiC)	1.2669(TaC) 2.3673(ZrC)	17.8084	14.6843
Theoretical CO_2 Actually Produced, %	72.87	84.03	0.36(B_4C) 0.26(SiC)	87.69(TaC) 46.93(ZrC)	0.95	6.72

[a] At controlled oxygen flow rate of 1 liter per min at 2800 F.

Table 3 — Oxidation Data on Dense Silicon Carbide Body

Specimen No.	Exposure Time, min	Exposure Temperature F	Results after Exposure
K1			Specimen in as received condition
K2	6	3100	No visual effect on specimen
K3	12	3100	Slight formation of thin-walled glass bubbles
K4	18	3100	Formation of glass bubbles over approximately ¼ exposed surface area
K5	2.5	2500	Slight formation of thin-walled glass bubbles
K6	2.5	3500	Slight formation of thin-walled glass bubbles
K7	10	3700	Formation of thin-walled glass bubbles over approximately ⅛ exposed surface area. Slight glaze on the rest of surface
K8	85	3700	Glass bubbles or glazing over entire surface which had been seriously attacked

Table 4 — Modulus of Rupture Test Data on Aluminum Modified, Dense Silicon Carbide Body[a]

Test Bar No.	Time, min	Temperature, F	Modulus of Rupture[b]	Remarks
1	—	80	32,100 psi	
2	—	80	30,500	
3	—	80	26,100	
4	—	80	25,600	
5	—	80	24,200	Off-center break
6	—	80	17,700	
7	3	1,500	17,000	
8	3	1,500	14,400	
9	60	1,500	17,900	
10	60	1,500	11,900	Off-center break
11	60	1,500	10,100	
12	60	1,500	8,500	Flaw in X-section
13	3	2,000	15,100	
14	3	2,000	9,500	
15	60	2,000	22,000	
16	60	2,000	15,200	
17	60	2,000	9,200	
18	60	2,000	8,300	Flaw in X-section
19	3	2,400	9,200	
20	3	2,400	6,300	
21	60	2,400	13,200	
22	60	2,400	12,200	
23	60	2,400	9,300	
24	60	2,400	9,200	
25	3	2,600	10,700	
26	3	2,600	7,600	
27	60	2,600	10,800	
28	60	2,600	7,400	
29	60	2,600	6,200	
30	60	2,600	3,500	Large flaw in X-section
31	3	2,700	10,400	
32	3	2,700	8,700	
33	60	2,700	17,200	Off-center break
34	60	2,700	12,000	
35	60	2,700	10,000	
36	60	2,700	9,300	

[a] Specimen size about 2½ × ¼ × ¼ in.
[b] Tested over a 1.9-in. span. Rate of loading — maximum load obtained in about 1.5 min.

Table 5 — Modulus of Rupture Test Data of Commercially Available Dense Silicon Carbide Body[a]

Test Temperature,[b] F	Test Results,[c] psi	Average of Results, psi
80 ± 5	38,300	36,100
	33,900	
1,500 ± 15	30,150	26,550
	22,500	
	27,000	
2,000 ± 25	33,500	27,530
	28,800	
	20,300	
2,500 ± 25	32,300	31,300
	30,300	
2,700 ± 25	13,900	13,300
	12,700	

[a] Specimen size approximately 3 × ½ × ¼ in.
[b] Specimens inserted directly into the furnace at the test temperature. All specimens were heat soaked for 1 hr at the test temperature before application of the load.
[c] Tested over a 2-in. span using a one point loading technique. Rate of loading about 34 lb/min.

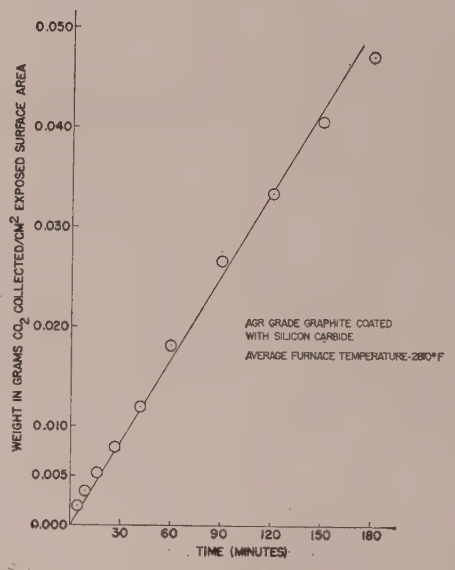


Fig. 2 — Weight CO_2 (grams/cm^2) surface area versus time

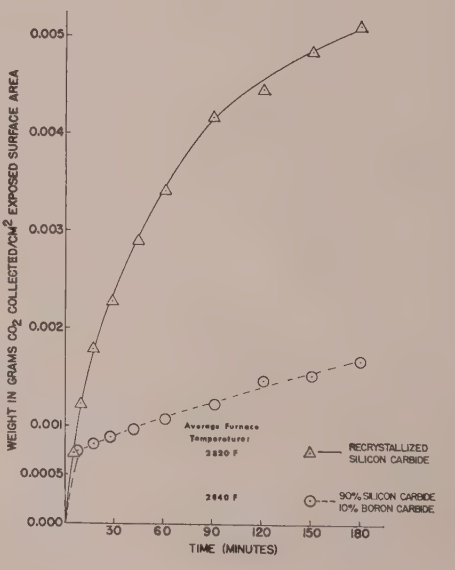


Fig. 3 — Weight CO_2 (grams/cm^2) surface area versus time

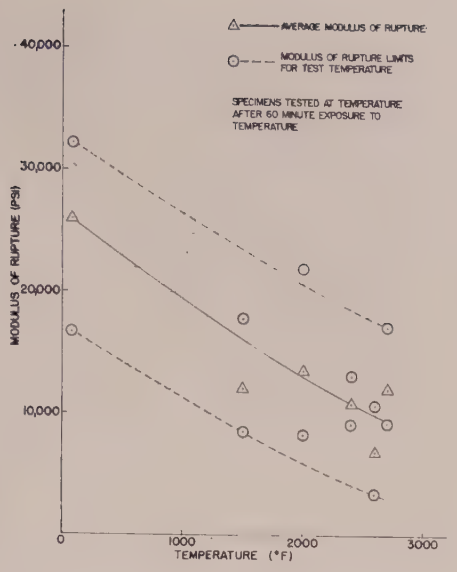


Fig. 4 — Modulus of rupture of modified dense silicon carbide body versus temperature

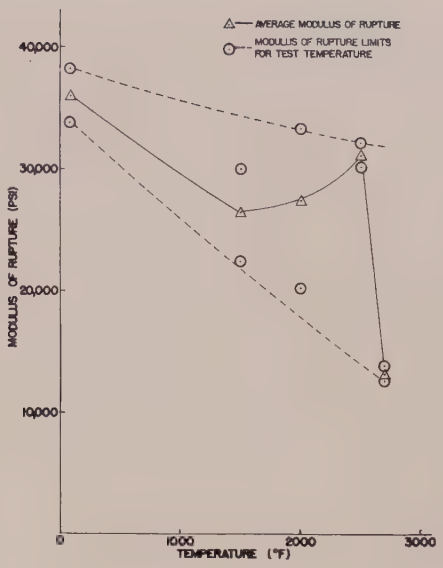


Fig. 5 — Modulus of rupture of dense silicon carbide body versus temperature

test samples in an oxy-acetylene furnace up to 3700 F were conducted on a 75% magnesia-25% zirconia body, a dense silicon carbide body, and a zirconium carbide body. The magnesia-zirconia body has been used as a funnel for transferring molten stainless steel. Visual examination of the test specimens showed little evidence of attack on the dense silicon carbide exposed for 10 min at 3700 F. Exposure for about 85 min resulted in severe oxidation of the body. Table 3 lists more complete visual data taken on the dense SiC specimens exposed. A substantial weight loss was noted after exposure of the MgO-ZrO_2 body for 1 hr at about 3450 F. A white coating was formed over the ZrC test specimens after 2–6 min exposure at temperatures of 3275–3450 F. The coating was observed to have poor adhesion to the body.

Considerable spread in modulus of rupture test results was obtained at temperatures to 2700 F for the aluminum modified, experimental, dense silicon carbide and to 2000 F for the commercially available dense silicon carbide. Test data for the modified dense silicon carbide for the 60-min exposure at temperature is plotted in Fig. 4. Table 4 compiles the test data collected for this body. Data for the commercially available dense silicon carbide is plotted in Fig. 5 and compiled in Table 5. Considerably lower modulus of rupture values were obtained for the ZrO_2 as shown in Table 6.

The hot-pressed 90/10 silicon carbide-boron carbide bodies exhibited considerably higher modulus of rupture strengths than the slip cast bodies. A hot-pressed body had an average modulus of rupture of 15,930 psi at 2700 F, whereas the best slip cast body had a modulus of rupture of 9390 psi at

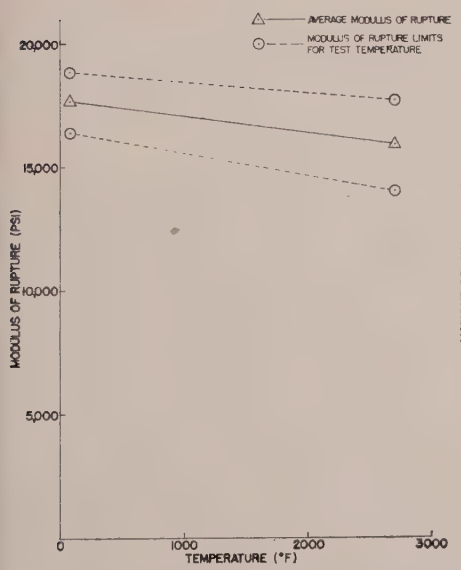


Fig. 6 — Modulus of rupture of 90/10 silicon carbide-boron carbide body (hot-pressed, fired 15 min at 4100 F) versus temperature

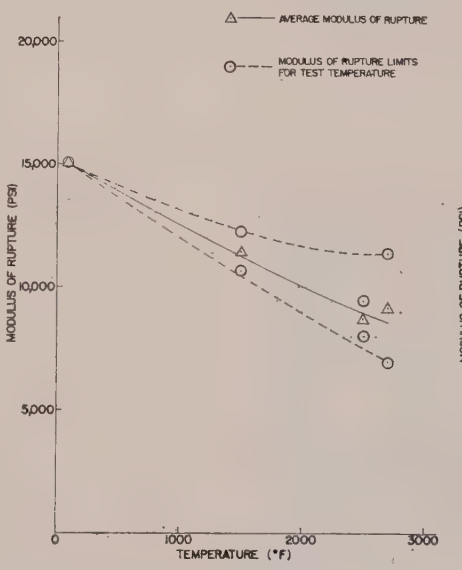


Fig. 7 — Modulus of rupture of 90/10 silicon carbide-boron carbide body (slip cast, fired 20 min at 3812 F) versus temperature

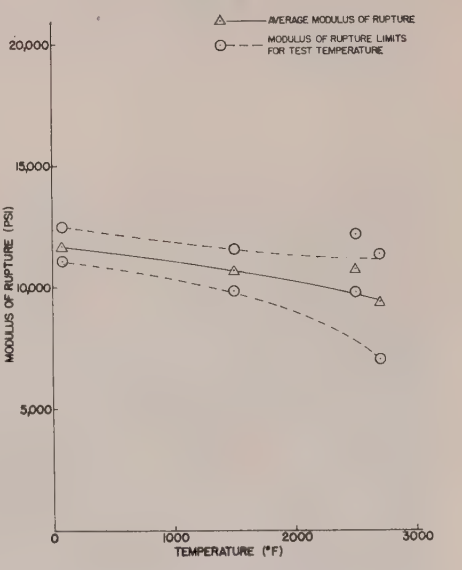


Fig. 8 — Modulus of rupture of 90/10 silicon carbide-boron carbide body (slip cast, fired 10 min at 4172 F) versus temperature

Table 6 — Modulus of Rupture Test Data on Zirconium Oxide Bodies[a]

Test Bar No.	Time, min	Temperature F	Modulus of Rupture, psi[b]
W-1129-1	—	80	4,200
W-1129-2	—	80	3,500
W-1129-3	—	80	3,500
W-1127-1	—	80	4,050
W-1127-2	—	80	3,500
W-1127-3	—	80	3,700
W-1128-1	—	80	2,400
W-1128-2	—	80	1,700
W11128-2	—	80	1,700
W-1129-4	3	1,500	Below 100[c]
W-1129-5	3	1,500	Below 100
W-1129-6	3	1,500	Below 100
W-1127-4	—	1,500	3,500
W-1127-5	—	1,500	3,700
W-1127-6	—	1,500	2,700
W-1128-4	—	1,500	2,200
W-1128-5	—	1,500	2,700
W-1128-6	—	1,500	2,800

[a] Specimen size approximately 4 × ½ × ½ in.
[b] Tested over a 3-in. span using a one point loading technique. Rate of loading — maximum load obtained in 1.5 min.
[c] With exception of the W-1129 body tested at 1500 F, all specimens tested at elevated temperature were inserted into the furnace at room temperature and heated to test temperature before breaking. Low values obtained may be due to thermal shock.

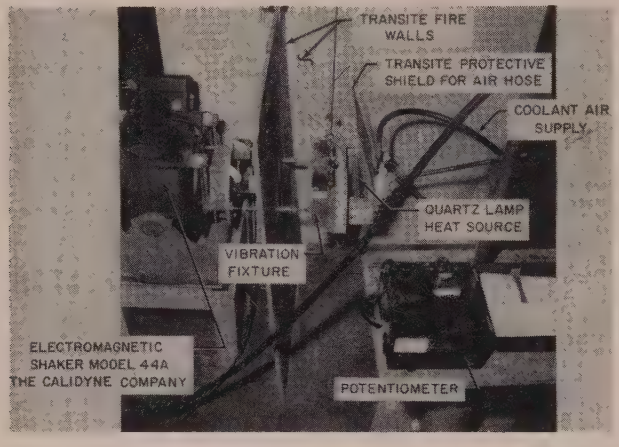


Fig. 9 — Test equipment for vibration of thermal insulants at high temperatures

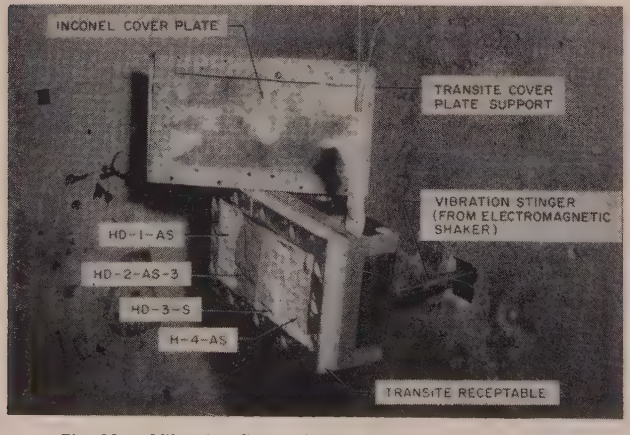


Fig. 10 — Vibration fixture before assembly of cover plate

that temperature. Detailed test results are given in Table 7 and plotted in Figs. 6–8. As may be seen, the silicon carbide-boron carbide bodies had an appreciably lower strength than the commercially available silicon carbide at temperatures below 2700 F.

Thermal Insulants — Vibration tests at temperatures of 1800–2000 F and water absorption tests were conducted on a number of low-density, silica and alumina-silica fibrous insulants.

1. Vibration at elevated temperatures.

Alumina-silica and silica fibrous insultants with a nominal density of 3–4 lb/ft^3 were vibrated at elevated temperature in three mutually perpendicular planes for 48 hr according to a predetermined schedule, at frequencies of 5–2000 cps with 0.06-in. double amplitude. These conditions are normal for some missile applications. Temperatures of about 1800–2000 F for most of the test were recorded on the hot side facing the heat source. No attempt was made to cool the other side of the insultant where temperatures in the range of 800–1000 F were recorded. Fig. 9 illustrates the test equipment used in our vibration studies. Fig. 10 shows the position of the insultant test specimens in the fixture.

After the 48-hr test it was observed that alumina-silica fibrous insulant, our number HD-2-AS-3, had the least physical deterioration. Alumina-silica fibrous insulant, our number HD-1-AS, exhibited the least shrinkage but was badly pitted by its impurities. Additional test specimens of the materials tested for 48-hr vibration at elevated

Table 7 — Modulus of Rupture Data of Various Silicon Carbide-Boron Carbide Bodies[a]

Specimen Composition, % by weight	Test Temperature, F[b]	Test Results, psi[c]	Average of Results, psi
90 SiC-10 B_4C	80 ± 5	16,450	17,680
		18,900	
Hot pressed. Fired 15 min 4,100 F	2,700 ± 25	17,800	15,530
		14,050	
90 SiC-10 B_4C	80 ± 5	15,100	15,090
Slip cast. Fired 20		15,080	
min 3,812 F	1,500 ± 15	10,670	11,490
		12,300	
	2,500 ± 25	9,450	8,750
		8,040	
	2,700 ± 25	11,400	9,190
		6,970	
90 SiC-10 B_4C	80 ± 5	11,350	11,650
Slip cast. Fired 10		11,200	
min 4,172 F		12,500	
		11,560	
	1,500 ± 15	11,580	10,710
		9,870	
		10,080	
		11,300	
	2,500 ± 25	4,300	10,763
		10,310	(excluding
		12,200	4,300 psi
		9,780	result)
	2,700 ± 25	11,380	9,390
		8,630	
		7,000	
		10,530	

[a] Specimens size approximately 3 × ¼ × ¼ in.
[b] Specimens soaked at temperature for 1 hr before testing.
[c] One point loading over a 2-in. span. Rate of loading about 34 lb/min.

temperature were vibration tested at elevated temperatures at an increased amplitude. An amplitude of 0.4-in. double amplitude was used. Such amplitude is greater than that expected for a useable structure. Due to the possible adverse effect of this amplitude on the test unit, the test was restricted to ½ hr in each plane, for a total of 1½ hr. All the specimens remained dimensionally unchanged except a high silica fibrous insultant. The 3 lb/ft^3 alumina-silica insulant HD-2-AS-3 showed the least change in density. A 12-hr vibration test at elevated temperatures was conducted on several alumina-silica fibrous insulants having densities of 12 lb/ft^3. Vibration tests were conducted on the higher density materials primarily to correlate vibration resistance with density. With the exception of some powdering, the materials appeared to withstand the vibration satisfactorily.

2. Water Absorption.

A water immersion test was conducted on all the insulant specimens evaluated. All the insulant test specimens absorbed water in the "as received" condition. The amount of water absorbed by the specimens having low "as received" water absorption values increased after vibration testing. None of the specimens retained water after oven drying. There was no evidence of hydrolysis or chemical reaction of water on any of the specimens.

Conclusions

Refractory Materials for Hot-Spot Areas —

1. Of the carbide materials tested the best oxidation resistance at 2800 F was exhibited by the refractory body consisting of 90% silicon carbide and 10% boron carbide (SiC-B_4C).

2. At about 3700 F the standard dense silicon carbide body showed little deterioration after 10-min exposure in any oxy-acetylene furnace. Specimens subjected to 85-min exposure at these temperatures deteriorated quite badly resulting in loss of about one-fourth of the original volume of the specimen.

3. An attempt to obtain by the addition of aluminum a commercially producible silicon carbide possessing a modulus of rupture approaching that of a hot-pressed laboratory body was not successful. The strength of the aluminum modified body was less than that of the commercially available dense SiC.

4. The graphite coated with SiC shows promise for applications at temperatures of at least 2800 F and possibly 3700 F for short periods. Oxidation tests would have to be conducted at 3700 F on the coated graphite. Coated graphite has a substantial weight advantage over SiC. The density of graphite is about 1.8 gram/cc. The density of commercial varieties of silicon carbide vary from about 2.5 to 3.1 gram/cc.

5. The dense silicon carbide bodies had considerably higher modulus of rupture strength than the silicon carbide-boron carbide bodies and the zirconia bodies at most of the test temperatures.

Thermal Insulants —

1. Based on the tests conducted it appears that an alumina-silica fibrous insulant (our code HD-2-AS-3) having a nominal density of 3 lb/ft^3 shows the most promise as an insulant for our application. Selection of an insulant will be based on tests conducted in this and previous studies and on the thermal conductivity of the insulants before and after environmental exposures.

2. The water absorption tests designed in the laboratory were used to compare the insulants for screening purposes only. It appears that the use of protective seals to prevent water absorption by fibrous insulants may be necessary.

Acknowledgment

Most of the work described in this report was sponsored by the U. S. Air Force. Their assistance is gratefully acknowledged.

DISCUSSION

Metal-Ceramic Composites Show Promise

— J. R. Tinklepaugh
Alfred University

MR. NOWAK is to be complimented on his excellent summary of a problem which is of real concern to many of us who are engaged in ceramic research. I do take issue with the use of the term "nonmetals" in the title rather than "ceramics." This would be excusable if the author were not a ceramic engineer.

The problem is not a new one, but has been anticipated for several years. Yet, the limited amount of data contained in Mr. Nowak's paper serves to illustrate the fact that very little sound, long-range development research has been done on the use of ceramic materials in high-stress, high-temperature applications. Too much of the sponsored research has been of exploratory type closely associated with a specific need in one type of application or another. This situation is not entirely the fault of the sponsors, but must also be shared by those of us who make the proposals and do the research.

For high-temperature high-stress applications involving a high degree of thermal shock resistance, beryllium oxide appears the most promising of the oxides from a purely technical viewpoint. Its scarcity and high cost may prohibit its use. Alumin oxide appears next in line and is already in widespread use. Of the carbides, silicon carbide seems to lead. Some compromises are necessary, however. The SiC developed in our laboratory having a cross-bending strength of 70,000 psi at 2500 F mentioned by Mr. Nowak was also more susceptible to thermal shock failure than silicon carbide in some other forms.

I have felt for some itme that should pure ceramics fail to meet some of the requirements, metal-ceramic composites may be found satisfactory. Our laboratory has investigated the fabrication of metal-ceramic laminates, metal fiber reinforced ceramics, metal mesh reinforced cermets and ceramics, and even ceramic reinforced alloys. Some excellent work has been done at the Armour Research Foundation and at Marquardt Aircraft on such composites. It has been demonstrated that a 9% volume addition of molybdenum fiber to alumina results in a composite that retains a cross-bending strength of the order of 15,000 psi following six thermal endurance cycles in which the pure alumina oxide samples were destroyed in one or two cycles. The field of metal-ceramic composites is still largely unexplored.

the first year

JET AGE . .

F. W. Kolk
American Airlines, Inc.

THE AMERICAN operating industry has now completed the first year of operation of some of the advanced equipment characterized popularly as "jets." The time seems appropriate to recount some of our experiences during this year and to assess the progress against our initial expectations in the planning stage of this equipment. During this period, my company has been responsible for a very large part of the total operation of these new aircraft and we have progressed from the pre-operational period of wide-eyed enthusiasm to the present coping with grubby everyday problems. There have been moments of elation during this time, and we have faced stark tragedy; but no one connected with this business fails to carry a deep-seated sense of satisfaction and accomplishment when looking back over this period.

We have found the public acceptance of the new aircraft to be phenomenal. Not only have these aircraft been operating at extremely high load factors, but they have led to a tremendous stimulation in business. They have created travel where travel did not before exist. How long this will continue, no one knows; but it appears certain that we will not regress to the previous levels of business.

It is our task, as an industry, to improve ourselves. We must now take stock of our overall record to find out what we have done well and what we have done less than well. This is not in an attempt to assess blame or stigma, but rather to apply some Monday morning quarterbacking so that we are better prepared for next Saturday's game. Of necessity then, I must deal not with the portions of this huge task which have been done exceptionally well, but rather concentrate on the problems found to examine where our foresight was lacking and to suggest remedies. The industry has just taken a revolutionary step and it will be a long time before we will be again ready to take another such step. The next decade for the aircraft industry will be one of consolidation, of development, of refinement, of overcoming problems that were not at all obvious three years ago. There is plenty of room for such improvement in today's generation of aircraft. When we can no longer think of such improvements, it is time to progress toward a radically new type of machine, such as the supersonic transport. Today's aircraft will be with us, however, for a lengthy period and we had best understand them well so that we can use them to the utmost advantage.

Scheduling Problems

When we ordered 707 jets and Electra turboprops in 1955, it was believed that the turboprop would be available for service quite a bit sooner than the big jet. Inasmuch as this turboprop would be by far the most attractive aircraft in our stable at the time of its delivery, it was logical to expect introduction of service on the transcontinental routes. This dictated a range requirement sufficient to allow operation nonstop from New York to Los Angeles. Therefore, the fuel system, the take-off gross weight, and the airport characteristics were established with this capability in mind. At the same time, it was realized that the jet airplane, when it did appear, would go into service on these routes. Even at that time, it was anticipated that we could expect to re-engine the 707 sometime during its life and it was believed that the revenue from the passenger business on these long nonstops was substan-

of the

. . reflections

THE FIRST YEAR of jet airline operation has brought many problems — and satisfactions — to the industry. Here the author discusses some of the more serious problems:

1. Scheduling. American Airlines used the "Monte Carlo" method to calculate payloads and flight times.
2. Baggage handling. Almost nothing annoys a passenger more than long waits for baggage at the end of a flight. One approach to the problem is the baggage expediter system.
3. Mechanical shutdowns.
4. Runway length.
5. Noise. Noise suppressors have not been effective enough, from the standpoint of communities surroundings airports. Development of the turbofan engine offers some hope in this area.*

tially more profitable than that from the cargo business. I want to emphasize that this is what our data showed at that time. Due to the rapid development of our business, these conclusions are no longer true. However, in layout of both aircraft, an attempt was made to produce designs which would operate successfully, purely as passenger aircraft, nonstop from New York to San Francisco under the conditions then believed to be typical of such a flight. Needless to say, we were disappointed in the range characteristics we realized from the aircraft. Successive increases in the weight of the earlier. Electra aircraft make it rather doubtful whether a transcontinental operation could have been successfully accomplished at the time of the delivery of the first aircraft. This was not entirely the fault of the flying machine; the environmental conditions for this flight have changed radically since 1955. The introduction of positive control of air space occupancy has served to lengthen the number of air miles required to cover the Great Circle ground distance between the city pairs and the problems of airway congestion and altitude separation are such as to limit further the aircraft's range abilities. Fig. 1 illustrates this effect by showing the departure routes from Idlewild Airport. The combination of these two limitations increases the practical range requirement by some 20% over the Great Circle optimum range flight, not counting the effects of head winds Had not the delivery of the 707 been improved, we could no doubt have modified the Electra to gain the necessary range.

In the case of the 707, we had deliberately attempted to create the range payload break at the transcontinental range, with the proper margins for operational requirements. With the aforementioned growth of the range requirement, plus the effects of such deficiencies as sound suppressors and a slight degree of overoptimism on the part of all hands in the matter of aircraft performance, we were in 1958 brought up short by the realization that we might not be able to carry a full load of passengers from New York to Los Angeles on the introduction of service. The problem can be posed a little like this — If we assume an average payload, an average wind aloft, an average fuel reserve, and an average ground temperature, we can make the flight handily with a fair amount of margin. If, however, we assume a full load of passengers, not quite the worst wind on record, optimum altitude blocked by other traffic, and the kind of fuel reserve you must carry when it is raining in Los Angeles, the sum of the parts becomes somewhat greater than the whole and the passenger load tends to dwindle to insignificance. Also, in operating a jet aircraft, such things as exact choice of flight altitude become quite important to the resulting flight time and fuel requirement. This contrast with the rather wide ability of the piston-engined types to change altitude without affecting fuel requirements to a very large degree. Some reasonably consistent allowance for off-altitude operation must be built into any scheme of calculating payloads and flight times.

It is possible to get data on the frequency distributions of wind, temperature, occupancy of altitude, routings, fuel reserves, and the like, and it is possible to compute a flight plan for every combination of these factors. The number of combina-

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 9, 1959.

Fig. 1 — Departure routes from Idlewild Airport, New York

tions possible, however, reaches astronomical numbers and we needed to invent some calculation scheme which would combine the relative likelihood or probability that certain combinations of circumstances would occur. The method we adopted is well-known in certain types of operational research studies and is called the "Monte Carlo" method. This is how it works:

To appreciate the complexity of the payload-block time problem, a block outline of the basic approach is shown in Table 1. The table lists eight items which are of significance in the conduct of the flight. Each one of these items is subject to almost infinite day-to-day variation and the frequency that certain values appear can be expressed by probability curves.

The distribution of any item can be expressed in the form of several sets of unique values, each of which is equally likely to occur. The number of individual values used to express this distribution is the degree of freedom of the variable. Thus, the wind is expressed as 10 equally likely values, the average of which represents the average wind.

It is obvious that the product of all the degrees of freedom represents the total number of equally likely environments in which flights could be conducted. Presumably, if we would calculate the flight plan for each combination, we could use the results to make a smooth probability curve of the actual operation.

There are two problems with such a procedure. First, the total number of possibilities is 5.29×10^9. The other is that for a given flight, only two altitudes can be used — one for each half of the flight. In choosing these altitudes, the best combination will be used; but the best must be based on the best available in light of airway congestion.

In order to simplify this mess down to a practical problem, we resort to a sampling technique. This technique can be compared to a roulette wheel, which has many numbers each of which is equally likely to show. We pick a value of each variable by giving the available values index numbers and then spin the wheel to get the one to use. In practice, the spinning of a roulette wheel or the flipping of a coin is rather time consuming, so a table of random numbers is used to generate the index digits.

If a sufficiently large number of situations are developed in this manner, the results of operation analysis under these conditions are quite representative of an actual operation. Too small a number of situations produce erroneous results, just as the analysis of too small a number of actual flights may be misleading; but reliable data can be obtained from far less than the 5.29×10^9 possibilities. We have found that 30 cases or situations are adequate for most work.

In effect, we duplicated the complete environment and management of each of 30 flights by mathematics. The results were plotted as probability curves of flight time resulting from 30 situations, exactly as we would have plotted actual flight times. The appropriate schedule times from the probability curves were used as the basis of published schedules. A corresponding plot of payload probabilities was used to determine the committed load. I must emphasize at this point, that committed load and schedule time are not mutually exclusive; that is, one can, within some limitations, increase the committed load at the expense of schedule time or vice versa. Our choice of a schedule was based upon a compromise which favored speed rather than payload in the belief that the basic commodity we have for sale is speed. The study would show that circumstances could conspire to prevent the completion of a nonstop flight but that this combina- was sufficiently unlikely to warrant any passenger restrictions in the early operation.

How did we make out in our predictions? It would seem from Fig. 2 that on the average, we pretty well anticipated the general characteristics of the operation. On certain routes, we missed the block time somewhat; usually because the estimates for such things as ground time and en-route delays, which were fed into the calculation, turned out to be incorrect. Apparently, we were somewhat better on the payload estimation because we have been able to carry the committed payloads on all routes, but have been unable to subsequently increase these loads beyond the original limits. We feel that we have developed a very valuable, although cumbersome, tool for the operational investigation of critical flight legs. The complication of this approach is not warranted under conditions where no real trade exists between payload and flight time. Our Monte Carlo analysis is related to a paper jet operation

Table 1 — Monte Carlo Analysis Program

	Degrees of Freedom	Distribution	Controlled by Flight
Take-Off Surface Weather	30	Random from available records	No
En-Route Wind	10	Normal	No
En-Route Temperature	10	Normal	No
Route Distance	6	Estimated	No
Available Fuel	49	Blocked at random	Yes
Fuel Over Destination	3	Estimated by weather records	No
Payload Available	20	To produce 85% load factor	No
Delay	100	Negative exponential	No

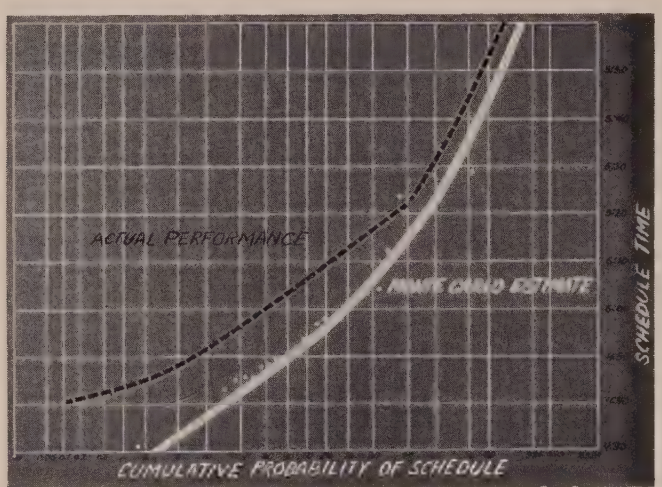


Fig. 2 — Schedule performance of B-707, Los Angeles to Idlewild, Spring, 1959

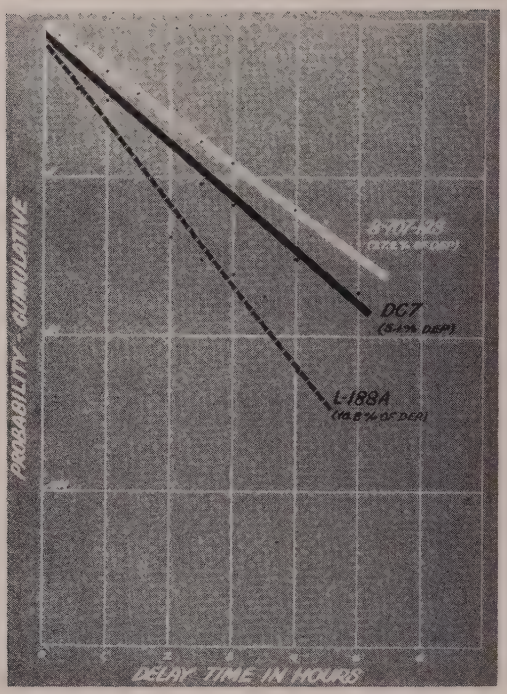


Fig. 3 — Mechanical delay, July, 1959

and, in fact, is a very complete paper jet operation — with provisions made to simulate such interferences as blocked altitudes due to traffic control, circuitous routings for the same reason, variable payloads, changing surface conditions at point to take-off, and the like. It was hammered out under conditions of duress immediately prior to the introduction of service and could and will be improved upon for future use.

The problem of speed versus payload compatibility is largely confined to the operation of the 707 aircraft. On the shorter routes, where the Electra is currently operating, few, if any, restrictions exist.

Baggage Expediters — In both the Electra and the 707, we have introduced machinery designed to increase the traffic effectiveness of the aircraft from the passenger's viewpoint. This is the baggage expediter system. The baggage expediter is applied to both airplanes according to a similar concept, although the detailed execution of the equipment is somewhat different in each case. The idea is not a new one and, in fact, had been experimented with in the early days of the DC-6. It consists of preloaded baggage containers which are inserted into the belly of the aircraft by means of a hoisting and traversing mechanism. The object of the system is to speed up the handling of the passenger's baggage, by shortening the time required after the landing of the aircraft to deliver the baggage to the passenger. The system is expensive, the containers are heavy and are subject to some wear and tear. We can say, however, that the results have been extremely successful — that they do achieve a significant reduction in the delivery time of the baggage and, quite unexpectedly, they have markedly reduced the amount of baggage damage and the resulting claims. The reason for the reduction in the damage is very simple. It is not necessary to manhandle an individual bag from a baggage cart into a baggage pit and stow it in the pit. The placing of the baggage into the container is done behind the check-in counter and the baggage is undisturbed until it is removed from the container at the delivery area of the airport of destination.

Mechanical Reliability

The ultimate ability of an airline to make money with an airplane and to provide satisfactory service to the public is tied up in the mechanical reliability of the vehicle. Before the turbine airplanes went into service, we felt that it was possible to achieve a much higher degree of reliability with turbine-powered equipment than had been the case in the past with piston-engine aircraft. The reason for this belief was the fact that the items contributing to mechanical irregularities were, to a very great extent, tied up in the powerplant, as opposed to the airframe, and that the essentially simpler thermodynamic and mechanical processes of the gas-turbine engine should lead to a more reliable operation. It has further been reasoned that the turbojet is superior to the turboprop in that it is the simplest basic propulsion concept and that the turboprop merely adds complication to the machinery present in the turbojet. Experience has, to a large measure, supported this view. The in-flight reliability of the turbojet engine has turned out to be extremely good by any standard and the in-flight reliability of the turboprop has been somewhat poorer than the turbojet. An analysis of the causes for in-flight shutdowns of the various powerplants indicates that the power section of the turboprop is quite the equivalent to the turbojet in operation. Such things as propeller control systems, gearboxes, and the like have been involved in a fairly high portion of the in-flight turboprop shutdowns. It would not be fair to say that this is the final conclusion — but it is true that, mechanically, the turboprop has been somewhat inferior to the turbojet from the standpoint of schedule reliability. On the other hand, the turboprop now being operated appears to have a higher rate of buildup of time between overhauls than the turbojet. But again, it is unfair to conclude that this is inherent in such engines but is rather the result of certain detail design problems of certain portions of the engine. One conclusion is completely supported by our experience. It is possible, in these engines, to locate the chronic source of trouble — the weak link in the chain — and it is relatively easy to design a fix for the particular part involved and to install it in the engine. In this regard, all turbine engines are infinitely superior to piston engines. The ultimate life of the turbine engines will extend to rather long periods between overhaul, once the critical parts have been brought under control.

The ingestion of foreign objects has not been a critical problem in the operation of these engines. There has been relatively little foreign object damage in either the jet or the turboprop. Our greatest vulnerability to such damage appears to be collision with sea gulls during take-off and climb.

It is possible to rate the mechanical excellence of an aircraft by the incidence of mechanical delays and by the length of these delays when they occur. Delays following a negative exponential distribu-

tion are essentially of random length and the probability curves of the delay performance are quite enlightening. Such data for the operation of the 707, L-188, and the DC-7 for the month of July, 1959, are shown in Fig. 3. These delays include all time chargeable to maintenance. This curve is based upon the probability of any flight being delayed. These records are accumulated on new airplanes which are well up on their learning curves and which contain certain chronic mechanical problems which are in course of correction. Examination of the delay causes shows that on the turbine equipment, a very small portion of the delays result from powerplant malfunction as opposed to a very large portion of the delay in the case of the piston equipment. The delays in the turbine aircraft at present are due to bugs in the aircraft system which can and will be corrected, and to lack of experience in trouble-shooting and maintenance procedures on a new and complicated piece of machinery. Both the systems and our knowledge of them will improve rapidly so that the aircraft system delays should be sharply curtailed. It can be seen, however, that the 707 does break down far more often than the Electra. Part of this is due to the fact that the Electra is a much simpler airplane than the jet and also to the exceptional effort made by the manufacturers of the Electra to design the utmost of reliability and maintainability into the airframe. Considering that the Electra was nothing more than a "sales brochure" four years ago, this performance is outstanding.

Operating Problems

We have had a number of operating problems develop with the jet airplane. At the same time, we have been virtually devoid of such problems with the turboprop. I suppose the reason for this should have been obvious to us, because although the turboprop uses a little different means to generate its power, it is still a propeller-driven aircraft. We all know how to build a good propeller aircraft, or at least we ought to after all these years. There are no unusual flight characteristics about the Electra which cause it to be treated differently from other aircraft. In the case of the jet aircraft, a number of new problems arose, which are peculiar to jets. Circumstances have forced us to operate the jet airplanes from runways which are barely sufficient in length to accommodate the required weight.

The first problem, which presented itself and which was realized by most operators fairly early in the game, was that the certified landing distances for intended destination of jet aircraft were unconservatively short for airline operation. This was not thought to be a problem in the initial phases of the operation and, indeed, it has not been a problem up until now because, in the initial long-haul operation, the take-off distance governed the field length and plenty of excess runway has always been available for landing. Operational evaluation of the 707, however, has indicated that the practical landing runway is some 15–20% in excess of the certificated values. This characteristic appears to be applicable to all jet transports of this general type and not specifically a 707 problem. Because of the lack of slipstream over the wings, requiring a higher approach speed than straight-wing craft and the relatively ineffective thrust reversers, the margins present in the certified landing maneuver and the standard 60% rule are insufficient for conservative, all-weather operation. The landing maneuver prescribed in the applicable regulations has permitted a technique of testing that has, on a number of occasions, caused major structural damage to aircraft when such a maneuver was actually flown. There exists an urgent need for a revision to the landing requirement so that realistic field lengths are produced by the certification of jet aircraft and, at the same time, preclude maneuvers during demonstration which jeopardize the airplane's structure. Such limitations must include a limitation on sink rate at the threshold point to such a value as would be chosen by pilots in actual service, so that the airborne portion of the requirement is based on a normal approach rather than an abnormal approach. Strict limitation of the vertical velocity at touchdown should also be prescribed and the approach speed must be limited to such a value that a pull-out from the threshold can be made with minimum altitude loss and without danger of falling below a minimum safe flight speed. Any regulation prescribing landing runway length must apply equally to piston aircraft, turboprop, and turbojets. We cannot have separate regulations for different powerplant types. In the meantime, we would suggest limiting the field length requirements for modern jet aircraft to the following empirical value:

$$D = \frac{V^2}{2.9} + 390 \qquad (1)$$

where:

D = Runway required, ft
V = Approach speed over threshold, knots

V should be at least a speed at which a positive gradient of climb is available with take-off thrust on all engines in the landing configuration.

The problems involved in take-off have been many and varied. All of the turbine aircraft have been licensed under regulations which require the full accounting of temperatures on all elements of the take-off. Evidence we have now indicates that the licensing of the 707 and Electra succeed in this aim admirably. In the 707 especially, the take-offs are extremely reproducible and the tightness of fit is unaffected by temperature, altitude, or runway length when the airplane is critically loaded. Take-off from a relatively short runway on a hot day at high altitude is no more critical than take-off from a long runway at sea level on a cold day, if the aircraft is operated at the limiting gross weight allowed by the airport analysis. Unlike previous aircraft, however, the jet aircraft *is* critical to all-engine take-off. When I say, "critical to all-engine take-off," I do not mean that the average performance on all engines is inferior to the performance with the loss of a critical engine at the critical point on the runway. But rather, that the overall hazard of the take-off is dominated by the all-engine condition. There are two reasons for this:

1. The all-engine take-off is closer to an engine-out take-off, in the case of the jet, than it has been with the piston types. That is, the performance degeneration due to engine cut is less in proportion than with piston types.

2. With the higher reliability of the powerplants

in jet aircraft, the chances of having an engine fail during the take-off are considerably reduced compared to our previous experience with piston-engine aircraft.

Any difference between the actual and the assumed conditions of the take-off, therefore, deteriorate the normal take-off performance to a point where some concern has been felt during some of the operations. One of the principle problems that has occurred is in the treatment of the wind component during take-off. The present regulations require that 50% of the head wind component be discounted and 50% of the tail wind component be added to the actual reported wind component for the take-off; but no stipulation is made as to the effects of a cross-component. Certain take-offs have appeared to be marginal under nominal zero wind conditions, and under cross-wind conditions where a zero head wind component has been assumed. In the case of the cross-wind, considerable lateral control must be used in the take-off flare, producing a good bit of spoiler action on the high wing. This has the effect of making the aircraft sticky as it breaks ground and results in a reduction in clearance over the end of the runway. Data on the performance deterioration incurred by this phenomena are not available to us and conservative action is required to account for it.

Another effect we have noted results from the turbulence or gustiness of the atmosphere. This gustiness affects not only the response of the aircraft, but also the wind direction and velocity data supplied to the pilot for his take-off computations. We have studied this gustiness phenomena enough to establish its general characteristics by making simultaneous readings of wind direction and velocity at short time intervals. These readings were analyzed by establishing a mean wind direction and velocity for a set and subtracting this vector from the instantaneous readings obtained:

where:

V_M = Mean wind vector
V_I = Instantaneous wind vector
V_G = Gust vector
$V_G = V_I - V_M$

Analysis of this data indicates that in a zero head wind component condition (either by a reported calm or a 90-deg cross-wind), a tail wind factor of 3 knots must be applied to the take-off airport analysis to determine allowable gross weights and to preserve the margins contemplated by the regulation (Fig. 4). This has led to the application, on American Airlines, of what is known as the rule of three (3 knot built-in tail wind component for take-off to our entire operation of jet aircraft). The application of this rule automatically gives us some 300 ft extra runway available for accelerate-stop and also gives us approximately 40 ft additional clearance on an all-engine take-off at the end of the pavement. We have found that day-to-day operation will use a good bit of this extra margin. The application of this rule is especially burdensome in conditions approaching a calm. Through the application of the 150% downwind rule, it serves to off load the airplane under conditions of 90-deg cross-wind or still air selectively, as opposed to the conditions of the usual 5–15 knot head wind condition. It gives us some protection against the existence of an overrunning, down wind component existing in the air above the runway surface, particularly at sea breeze airports during the time period when the wind is shifting from a land breeze to a sea breeze or vice versa. It may not be required for aircraft licensed under SR 422B (which requires a 15% margin on the all-engine take-off distance), but it certainly is on an SR-422 aircraft, which does not have all-engine requirement per se.

Fairly early in the operation of the 707, a take-off was made in conditions of quite heavy slush. The take-off performance of the airplane was extremely marginal under these conditions and, in addition, the force of the slush, thrown from a main gear against the fuselage resulted in some structural damage. The condition which leads to dangerous slush can be anticipated by meteorology—because a fairly heavy accretion of dry snow must be present,

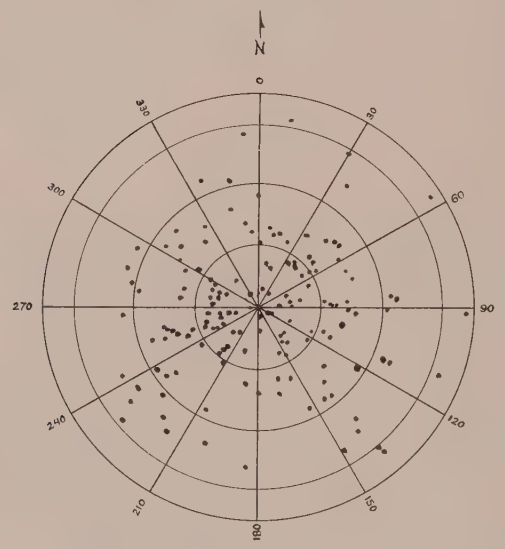


Fig. 4 — Surface gustiness vectors

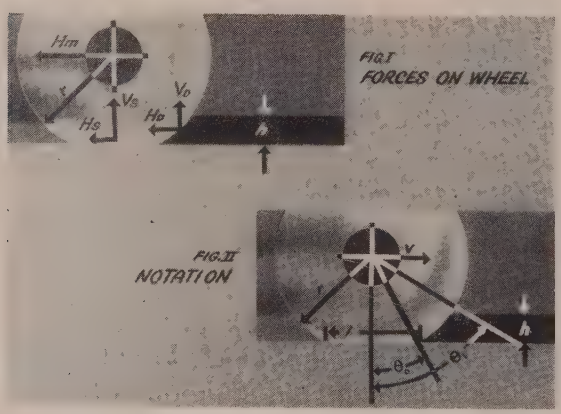


Fig. 5 — Wheel running through standing water

followed by an increase of temperature up to or slightly above freezing, with heavy rainfall so that the rainwater soaks into the snow accumulation. The result is a saturated mass of water and snow which has all of the essential hydraulic characteristics of water as far as interference with the aircraft is concerned and yet, which is retained in place by the suspension of snow so that it cannot run off. In this way, a homogeneous layer of fluid sticks on a runway to a constant depth, whereas free water would collect in puddles and rapidly run off. Any accumulation of slush of this type over half an inch in depth can be extremely hazardous to the take-off operation of jet airplanes, unless considerable excess runway exists relative to the required length for the weight to which the aircraft is loaded. An elementary treatment of the effect of this type of slush on take-off acceleration is contained in the Appendix.

The concept of a wheel running on the ground in the presence of a layer of fluid is a general one which is applicable to several interesting situations. (See Fig. 5.) When such a wheel intercepts the fluid, a part of the fluid layer is displayed and, by simple momentum concepts, the extent to which it remains displaced (does not re-expand and affect the retreating side of the wheel) represents a force which has both a drag and a buoyancy component. If the buoyancy force is high enough, it will support the tire without its contacting the "hard bottom" and when this condition occurs, "planing" takes place. There is a critical speed during a take-off in water or slush, when the drag buildup is at maximum; it corresponds to the condition when planning just starts to take place. In extreme cases, the resistance at this speed may get to be as high as the accelerating force and no further acceleration is possible. This phenomenon is similar to the "hump speed" of flying boats.

Typical effect of slush on take-off run is shown in Fig. 6.

Noise

The problem of community noise has been and is one of the most difficult problems facing us today with the jet aircraft. At the time the major decisions had to be made in the current new aircraft, a very considerable body of theory and demonstrated practice existed on the noise characteristics of propellers; very little was known about the basic mechanism of noise generation from jets. Thus, the gear ratio of the engine, the propeller diameter, and the number of blades could be chosen for the turboprop with due consideration of community noise. It was obvious that the Electra would be a quiet airplane on climb outs and approaches. So it has proved.

In the case of the jets, all that was known prior to the middle of 1955 was that the overall noise output varies as about the eight power of the jet velocity and the square of the diameter. From this knowledge we have consistently supported the development of engine cycles which would have low specific thrust. Such engines were not available in a proved form at that time. I believe that it was in June, 1955, that Greatrex presented his work on corrugated nozzles and their effect on sound radiation. The early test data looked good enough to warrant the introduction of jet aircraft into commercial service and orders were placed for the procurement of such aircraft.

When the jet aircraft were ordered, we in the industry realized that we could not operate an unsuppressed jet transport regardless of its economical attractiveness. The industry was determined to develop a satisfactory configuration of suppressor in order to make the conditions around the airport at least tolerable. By the same token, the industry was prepared to spend money, weight, and efficiency loss to achieve this result. We have more than spent our budget of these and have obtained an airplane which is about as noisy as the conventional propeller aircraft. In addition to the physical investment in the sound suppressor, we have been asked to invest in anti-noise flight patterns. Leaders in the communities affected by this noise have demanded that we climb out from the airports as steeply as possible, extracting maximum climb performance from the aircraft in order to alleviate the noise nuisance. We have complied with this desire insofar as prudence dictates, but we have been extremely reluctant to accept the responsibilities of extremely low climb-out speeds in view of the relatively poor performance and handling characteristics of jet aircraft as they approach stalling speed.

Quite unexpectedly, it appears that the major source of noise complaints have been on approaches rather than take-offs. Sufficient thrust must be used in the approach configuration at fairly low altitudes so that a considerable noise nuisance exists. This fact was not at once appreciated by the operators and the FAA; problems of traffic congestion led to traffic patterns which took the aircraft over populated areas for a considerable distance during approaches. These patterns are being revised to avoid as much excessive low-altitude maneuvering as possible.

What have the suppressors really done for us? In terms of the overall decibel level, they have produced startling results. Subjectively, however, they have served to suppress low-frequency noises relative to high-frequency noises. Having suppressed the low-frequency noises and shifted some of them into a high-frequency band, it is now generally accepted that the high-frequency components are more annoying than the low-frequency components of the same decibel level. We have, thus, been ac-

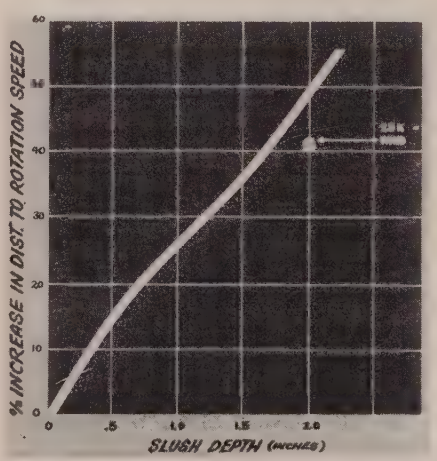


Fig. 6 — Increase in distance-to-rotation speed versus slush depth (B-707-123 of 247,000 lb, still air, sea level, 32 F)

cused of not accomplishing anything with the sound suppressor. In light of the manifest weight and performance penalties of the suppressor, this is very discouraging indeed. It may actually be possible to prove that the suppressor is a useless piece of hardware, although I do not believe we should be so pessimistic as to vigorously support this point of view. Certainly, the early suppressors were far from optimum and the purchasers of the early aircraft are stuck with them. All in all, the whole subject of noise suppressors has been disappointing. As one wag in the Pacific Northwest put it, "the aircraft seems to be equipped with a noise reverser/thrust suppressor."

Where Do We Go from Here?

From the foregoing discussion, it is obvious that the two airplanes American has introduced into service are behaving quite differently. Although both aircraft have a high degree of aceptance from the traveling public, the Electra seems to be fairly well devoid of operational problems, while the 707 has had many bothersome and, sometimes, serious problems. These result from the fact that the 707 in its intended service is critical to range payload, airport length, and community noise. A fundamental approach must be taken to correct these deficiencies. It would appear that a better miles/lb must be obtained through lower specific fuel consumptions, shorter landing field lengths must be obtained by lower stall speeds, and better take-off characteristics must be obtained by the use of more thrust in combination with the lower stall speeds. At the same time, an inherently quieter means of propulsion should be used to relieve the community noise nuisance, both in take-off climb out and approach. Furthermore, we owe it to our neighbors around the airports to achieve this desirable goal as soon as possible.

In my company, we have made some fundamental plans to achieve this end, including the modification, starting in late 1960, of all of our existing jet fleet. This modification consists of three fundamental elements:

1. The retrofitting of a leading edge glove to improve the high-speed capability of the aircraft.

2. Fitting of additional leading edge flaps to improve the low-speed characteristics of the aircraft—both from the viewpoint of stall and buffet speeds. These extra lead edge flaps should be of substantial help in coping with our runway length and climb out problems.

3. Conversion of the engines to a turbofan configuration. The turbofan, as has been pointed out on numerous occasions reduces the specific fuel consumption, increases the available take-off thrust and, by virtue of a sharp reduction in the jet velocity, considerably reduces the noise emanation. Each one of our major problems with the jet airplanes seems to be affected and alleviated by the installation of these engines. As a carry-on of this program, all additional airplanes of the jet type now on order by American will be powered by this or a similar competing engine. All aircraft are to be fitted with the best high-lift device compatible with the existing airframes.

Many people have asked what the overall practical effect of the turbofan engine is when it is applied to a jet aircraft. To answer these questions I have prepared, from information currently available, the pertinent characteristics of our Boeing 707–123 be-

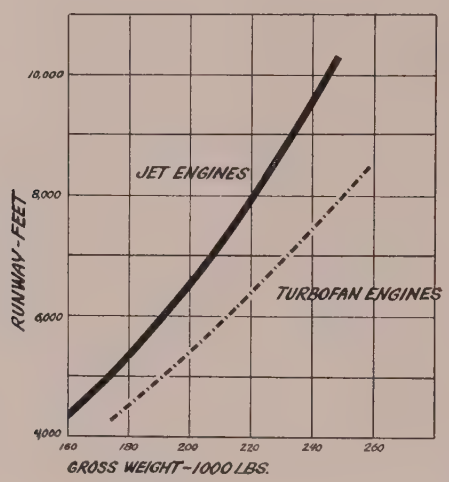


Fig. 7 — Take-off runway length improvement, B-707-123

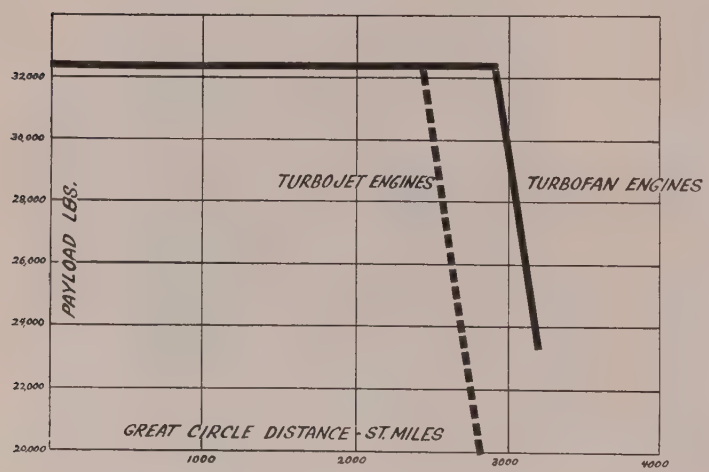


Fig. 8 — Range improvement, B-707-123

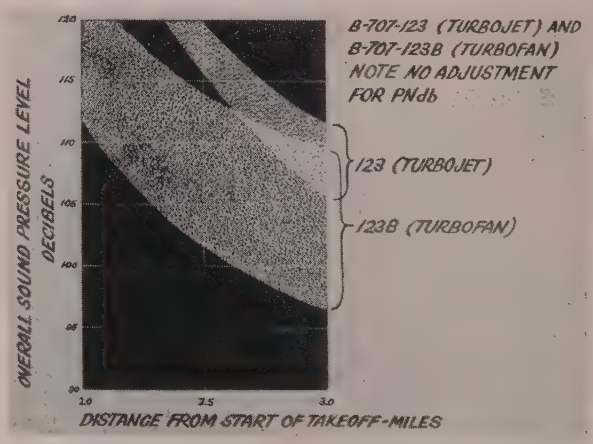


Fig. 9 — Estimated range of improvement in community noise level

fore and after its conversion to turbofan:

1. Take-off runway length required (Fig. 7). It can be seen that a very considerable improvement is herein forthcoming.

2. Range-payload (Fig. 8). Again it is seen that a very considerable improvement is to be had, sufficient so that operational margins to cope with devious routings, off-altitude operation, and the like are vastly improved.

3. Anticipated community noise level (Fig. 9). It is seen that the combination of lower noise output and higher altitudes are very productive in relieving the decibels. And note too, that because of the lack of high frequency shift, these are real, perceptible decibels.

4. Fig. 10 shows the cruising speed improvement.

In conclusion, jet transport has proved to be a remarkable tool to produce attractive transportation. It has had some problems and we are correcting these problems as expeditiously as possible.

APPENDIX

Resistance of Tire Rolling in Standing Water

Interference with normal rolling resistance, due to a layer of standing water or slush, adversely affects the acceleration of aircraft.

Scant measured data on the magnitude of this effect exists; nor is there likely to ever be any large volume of test data on this effect from which empirical conclusions can be drawn. It is necessary, therefore, to have a simple theory for the correlation of known data as it becomes available.

A first order theory is presented herein which appears to be useful for the purpose cited.

Forces in Action. (See Fig. 6.) — Consider a tire of radius r and width w inflated to internal pressure p. This tire is rolling at a speed v and encounters a layer of water of height h. The result is a set of forces as follows:

V_S = Vertical reaction of tire footprint area

$$= w\ 1\ p$$
$$= w\ p\ 2r\ \sin\theta_0 \qquad (1)$$

H_S = conventional rolling resistance term

$$= \mu\ V_S \qquad (2)$$

The forces V_D and H_D are dynamic in nature and stem from the fact that the intercepted area of the tire is subject to the stagnation pressure in the standing water. Actually, we are seeking the horizontal and vertical reactions of a single force:

$$V_D = \int_0^{V_D} dV_D = \int_{\theta_0}^{\theta_1} \tfrac{1}{2}\rho v^2 wr\cos\theta\, d\theta$$
$$= \tfrac{1}{2}\rho v^2 wr\,[\sin\theta_1 - \theta_0] \qquad (3)$$

$$H_D = \int_0^{H_D} dH_D = \int_{\theta_0}^{\theta_1} \tfrac{1}{2}\rho v^2 wr\cos\theta\, d\theta$$
$$\tfrac{1}{2}\rho v^2 wr[\cos\theta_1 - \cos\theta_0]$$
$$\tfrac{1}{2}\rho v^2 wr\left(\frac{h}{r}\right) \qquad (4)$$

A third dynamic force is exerted by spray reimpinging on a portion of the vehicle. A wheel displaces sideways, a mass of water per unit time:

$$m = \rho whv \qquad (5)$$

If the aircraft intercepts a portion *Km of this* displaced water, a drag is produced:

$$H_m = Kmv$$
$$= Kv^2\rho wh$$
$$= Kv^2\rho wr\left(\frac{h}{r}\right) \qquad (6)$$

For performance investigation it is necessary to know the effective rolling coefficient of friction, which is the ratio of the wheel drag to the wheel vertical load:

$$\mu^1 = \frac{H_S + H_D + H_m}{V_S + V_D}$$

$$\mu^1 = \frac{\mu wp2r\sin\theta_0 + \tfrac{1}{2}\rho v^2 wr\left(\frac{h}{r}\right) + Kv^2\rho wr\left(\frac{h}{r}\right)}{wp2r\sin\theta_0 + \tfrac{1}{2}\rho v^2 wr[\sin\theta_1 - \sin\theta_0]}$$

$$= \frac{\mu\sin\theta_0 + \rho\frac{v^2}{4p}\frac{h}{r}(1+2K)}{\sin\theta_0 + \rho\frac{v^2}{4p}[\sin\theta_1 - \sin\theta_0]} \qquad (7)$$

It is obvious that when $\theta_0 = 0$, the force V_S disappears. The significance of this situation is that the entire vertical load is supported by the dynamic pressure developed by water displacement. In this condition, "planing" occurs. In this regime:

$$\mu^1 = \left(\frac{h}{r}\right)\frac{(1+2K)}{\sin\theta_1} = \frac{(1-\cos\theta_1)\ (1+2K)}{\sin\theta_1} \qquad (8)$$

The following expression is approximately true:

$$\frac{1-\cos\theta_1}{\sin\theta_1} = \tan\left(\frac{\theta_1}{2}\right) \qquad (9)$$

It is to be noted that for a planing surface, the drag due to "lift" has an identical form. When planing occurs, there is another source of drag not considered here, namely that due to skin friction on the planing surface. This drag is not present in this case because the wheels are free to rotate. If they were restrained, as by brakes, some allowance would have to be made for this fact.

In order to evaluate the above expression, it is necessary to evaluate θ_1 and θ_2 for a given speed condition. This must be done by equating the total axle lift for all axles to the aircraft weight less wing lift. Such a relationship is a special one for any aircraft, depending on its geometry.

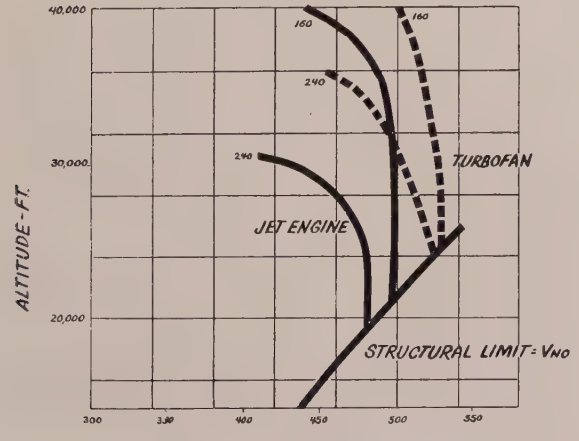


Fig. 10 — Cruising speed improvement

How Good Is Testing?

a correlation of customer, laboratory, and proving ground experience

W. A. McConnell
Ford Motor Co.

TEST ENGINEERS are increasingly concerned about how well their test routines and procedures reveal defects which the customer would encounter if he were to receive the undeveloped car. How much significance can be attached to a particular test result on a limited sample of a pre-production product? This concern is evident in several recent SAE papers on testing. To quote one reflecting common thinking:

> "Customer complaints and field service failures must be continually reviewed and analyzed to make sure no critical factors are overlooked — to establish the reliability of prototype vehicles as test samples — to check the accuracy with which our adopted sets of circumstances reflect what the customer actually will encounter."

In giving a brief report on the state of the testing art, we would like to draw a distinction between Performance Testing which is concerned with how *well* the product fulfills its intended function, and Endurance Testing, which is concerned with how *long* the product will continue to function as intended.

The primary problem in correlating customer experience with endurance tests involves simulating customer usage on the proving ground and in the laboratory in less time than the normal life span of the vehicle. It is necessary, to confirm the reliability of this simulation, to make comparisons of the failure distribution and frequency of occurrence on production vehicles in the field with similar data from the proving ground.

In the matter of performance tests, the aim is to establish operating conditions and performance criteria representative of customer requirements. But in evaluating and comparing the efficacy of various designs, where differences may be exceeding small, some other problems raise their ugly heads. How can we eliminate from the comparisons the effect of extraneous variables such as the driver and the weather? How accurate are the results, how significant are the data, and how many tests must we make to be confident of our conclusions?

Endurance Testing — Proving Ground

Development of a proving ground route to simulate average customer usage in less than normal time began four years ago with establishment of design specifications for the road construction at our newest Michigan Proving Ground.

For proving ground testing, we concluded that there were only three ways in which the endurance test process could be accelerated. One, the routes could be made arbitrarily more severe than any customer would experience, with sharper turns, steeper grades, and rougher surfaces, subjecting the vehicle and its components to higher than actual loads and stresses. Two, the routes could be made

THE METHOD of validating proving ground test routines as a sample of customer operations is described. Comparisons of field survey and proving ground failure statistics are analyzed to establish sample reliability and severity ratios. Alternative means of accelerating the test program are considered, with examples indicating the effect on correlation.

Use of proving ground routes as a link in establishing laboratory test to customer service correlation is treated. Procedures and instrumentation for developing load cycle information on the road, and programming random load cycles into laboratory test fixtures for simulated service life testing are discussed with some examples of applications to automotive components.

In the performance test area, several performance parameters representative of customer requirements are developed. Techniques of improving comparative evaluations by eliminating the effect of extraneous variables such as the driver and the weather are presented. Accuracy, sample size and confidence level of performance test data are revealed, with indications of the significance which can be attached to differences in comparative test data.*

with the worst features no worse than would be encountered in the field, but with much of the milder operation omitted, thus giving a higher frequency-of-occurrence of the highest load and stress cycles. Three, we could operate the vehicle more of the available time than the average customer is able to do.

In laying out the 4000-acre, 30-mile sample of the national highway system, we decided the vehicle should not be subjected to abnormally severe use, but should receive more frequent than normal application of the most severe conditions actually met in service. If the vehicle is considered as several thousand parts, each with its own fatigue life curve, the difficulties inherent in any other approach to an accelerated test schedule are apparent. If quicker failures are sought by arbitrarily increasing the vehicle loads, speeds, and stresses, nontypical failures are produced. However, if we delete the large number of low-stress cycles arising from milder operation and so achieve a more frequent application of higher stress conditions found in service, correlation with normal use should not be destroyed. We would hope to reduce only that part of the operation below the endurance limit, where no failures and little cumulative damage would occur.

Any acceleration of the test schedule, and round-the-clock driving, will necessarily result in some critical conditions being missed, notably ageing, corrosion, weathering, and soaking influences. These time-dependent factors must be investigated in other ways. We feel the validity of this approach is supported by some of the resulting data.

Fig. 1 illustrates one of the comparisons made four years ago between a hypothetical proving ground route we planned to build and a public highway course in southeastern Michigan we formerly used for test purposes. This public highway course had been mapped out, over a 10-year period, to include the most severe features available. The proving ground route contains the same mix of geometric features as the highway route, but they occur three times as often in a given mileage. Driving speeds are similar, although the more frequent stops result in a lower average speed for the proving ground. Similar occurrence measurements were made on a survey vehicle of vertical accelerations, suspension bottoming, and the like. The proving ground roads were designed to produce the *same* level of severity for all these disturbances as the public highways, but with *three times the concentration* as the southeastern Michigan route.

Comparison of proving ground failures with customer complaints have been handicapped in the past by the fact that most of the proving ground test effort has been on experimental, pre-production cars. Only a limited sample of production vehicles normally are tested over the proving ground roads to get a direct comparison of customer versus proving ground failures on identical vehicles.

Ideally, of course, there should be no correlation between prototype failures and field complaints if the development engineer has done his job well, since the whole purpose of proving ground testing is to find and correct mistakes before the customer finds them.

With abandonment of endurance test operations on the public highway and substitution of routes on

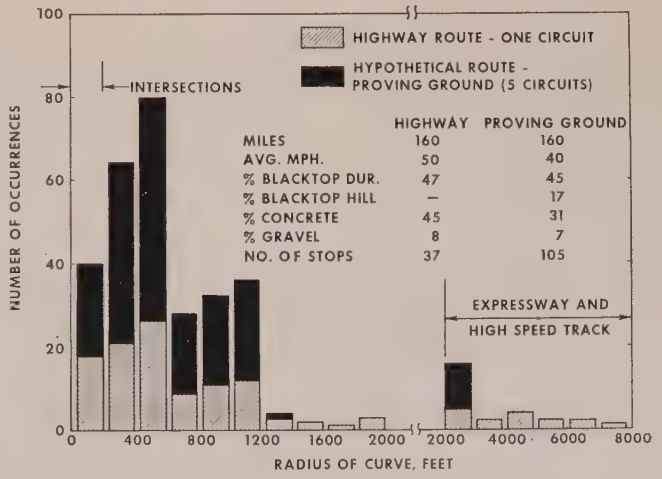


Fig. 1 — Comparison of public highway durability test route and tentative proving ground test route

* Paper presented at meeting of SAE Detroit Section, Dec. 7, 1959.

our new proving ground, it became desirable to test a large sample of production cars on the proving ground to replace the background of experience on previous production vehicles which had been accumulated over the years. This permitted a direct comparison of proving ground to field experience with a statistically significant fleet of test vehicles from the same model year.

Fig. 2 shows the number of failures, by uniform parts classification (UPC), which occurred on proving ground cars, compared with claims from the field. In general, *where there were customer problems, there were proving ground failures.* Only those UPC groupings with numerous customer complaints and high warranty costs are shown.

Several dozen cars were included in the proving ground sample. However, there were different numbers available for comparisons of production parts in the different UPC groupings. A more significant representation of the test data is to compare the number of failures in each UPC grouping as a per cent of the number of cars involved, as in Fig. 3, even though the exact number of vehicles generating field complaints is uncertain.

Expressed as a per cent, the proportion of proving ground failures is much higher. Note that there is a *20/1* ratio in the vertical scales. This means that if claims originate from 2% of the cars in the field, we can expect similar difficulties on 40% of the proving ground fleet.

The average mileage accumulated on the proving ground cars included in this comparison over the proving ground route was about 30,000 miles. However, field complaints were reported only for the 4000-mile warranty period. A better way of viewing the significance of the 20/1 ratio of percentages, therefore, is to consider that a vehicle tested for 30,000 miles on the proving ground will generate as many failures as 20 vehicles driven by our customers for 4000 miles each.

Expressed in terms of failures per car-mile, which is indicative of the severity of the route, we find slightly over two times as many failures per car-mile on the proving ground as there are warranty claims per car-mile from the customer. This is somewhat lower ratio than we might expect from a deliberate effort to create three times the frequency-of-occurrence of the severest loadings. By way of explanation, we note that proving ground procedure falls short of producing the expected proportion of fail-

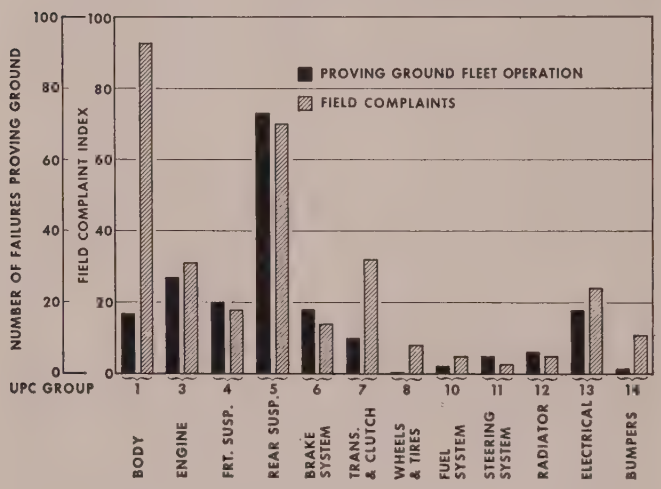


Fig. 2 — Number of failures during proving ground operation versus number of field complaints on similar items

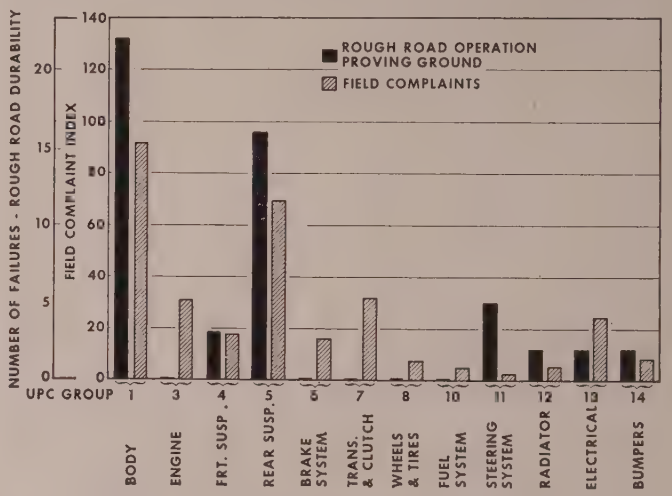


Fig. 4 — Number of failures during rough road operation versus number of field complaints

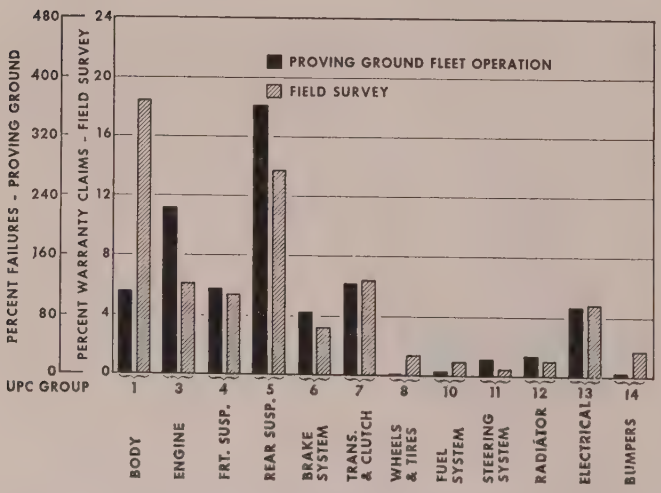


Fig. 3 — Comparison of per cent failures, proving ground operation versus field survey results

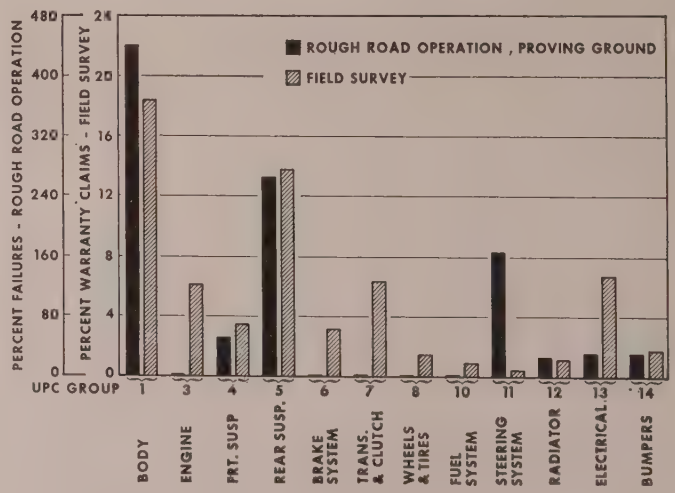


Fig. 5 — Comparison of per cent failures, proving ground rough road operation versus field survey results

ures in the body UPC groupings. Customer complaints in the body area are largely concerned with defective finishes, rusting of ornamental parts, and certain chronic functional disorders that include leaks, shakes, squeaks, and rattles which we normally report only once on proving ground tests.

If the calculation of relative severity is adjusted to include only the mechanical and structural type of customer failures in the body group, we find *2.5 times* as many failures per car-mile on the proving ground as occur in the hands of the customer. *This is the common rule of thumb we have used to relate proving ground to customer operation.*

Test cars have been operated on another route at our Dearborn Proving Ground, known as the rough road durability route. This route consists of about a 10-mile circuit, with a high proportion of Belgian block, cobblestone, rough gravel, and sharp turns.

Making a similar comparison of rough road route operation with field survey results, we find little correlation in many areas. This route, for example, does not develop engine and brake failures. However, body, suspension, and steering system defects are magnified (Fig. 4).

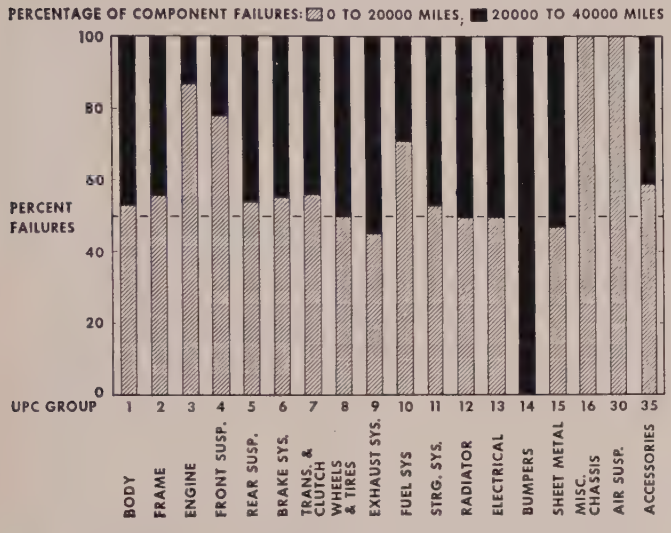


Fig. 6 — Distribution of failures during proving ground durability operation, 1957 and 1958 production and prototype cars

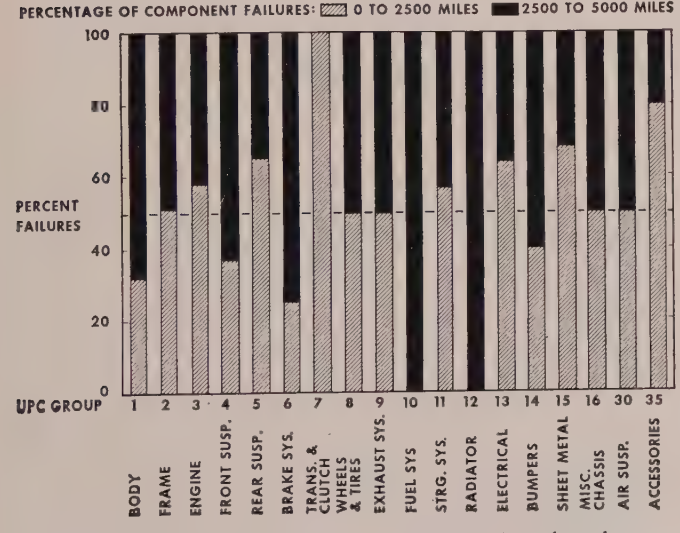


Fig. 7 — Distribution of failures during proving ground rough road operation, 1957 and 1958 production and prototype cars

In Fig. 5, this data is shown on a percentage basis to adjust the different number of cars participating. Figuring as we did before, there is again about a 20/1 ratio of per cent failures between *certain* of the UPC groups and field survey data in the matter of producing failures — but there are no failures at all in five of the groups. While body failures appear to correlate, the failures observed are not always typical of field complaints, consisting as they do of weld and metal failures where field claims are relatively low. This is also true in other areas. The high incidence of rear axle housing fractures on the rough road course is likewise disproportionate. Since the degree of magnification, or the severity ratio of the route, is not the same for different components, we can find no sound basis for relating this abnormally severe operation to customer use. While it is doubtless true that 5000 troublefree miles on the rough roads may assure no trouble on the highway *in selected areas,* it might also mean that some other parts are overdesigned. We suggest that this type of operation be used only to check progress, or relative merit of alternate designs, and that adequacy or inadequacy be established by operations more closely approaching normal use.

Figs. 6 and 7 are presented to show the distribution of failures with mileage. On proving ground durability operation about as many failures occur up to 20,000 miles as occur between 20,000 and 40,000 miles. Also, in very few instances do failures develop in the last 20,000 miles of a type different than those occuring in the first 20,000 miles. The exceptions are for the most part items where age and weathering is a factor. This suggests that there is little value in extending the durability operation beyond 20,000 miles. The purpose of this chart is to indicate that failures are nearly equally distributed throughout the mileage. This also has been evident on other durability tests of our own and competitive products, in which accurate maintenance cost records are maintained throughout a 20,000-mile period. The chief virtue to extended mileage is to assure development of typical failures on a limited sample, and thus increase its validity.

On the rough road or cobblestone route, failures are not as uniformly distributed with mileage, although again, typical failures are evident with few exceptions in 2500 miles.

Endurance Testing — Laboratory

Endurance testing in the laboratory is done with much the same approach as road testing. Applied loads must be realistic if failures are to be typical. The compressed time schedule is obtained by more frequent application of normal loads, and not by an arbitrary increase in operating severity. The major problem here, as on the proving ground, is in establishing some relationship between service life to the customer and the number and magnitude of the artificial load or duty cycles to be applied.

To this end much progress has been made in developing and using statistical instruments to collect background data from highway and proving ground routes. For any given operational route, the two major factors required to permit laboratory simulation are "how much" and "how many times." If we know how many times a vehicle component is subjected to the various levels of usage in its ex-

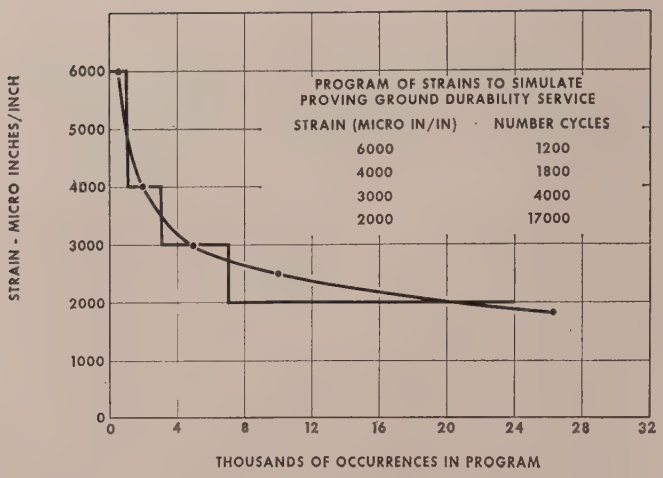


Fig. 8 — Strain versus frequency of occurrence on proving ground durability route — typical chassis component

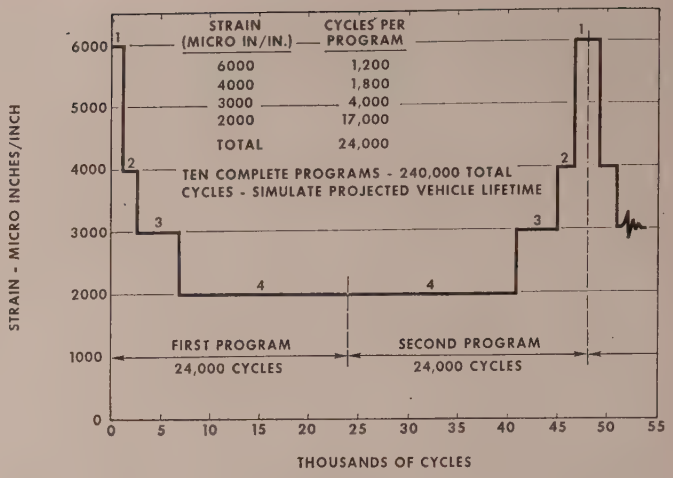


Fig. 9 — Laboratory programming to simulate proving ground durability route — typical chassis component

pected lifetime, we generally can set up a simulated test in the laboratory that will give realistic results.

There are many items that are of the "on-off" or "open-shut" variety which require only knowledge of "how many times" to set up laboratory tests. Most electrical switches or controls, doors, deck lids, and window and hood operating mechanisms fall into this category. Instrumentation to accumulate this type of data is relatively straight forward. An electric switch located at the item under observation is positioned to actuate a counter when the item is operated. It is necessary only to operate the vehicle in normal customer fashion to obtain good operational data. Generally, many items are observed on the same vehicle, each having its own switch and counter.

Unfortunately, the majority of vehicle components are subjected to varying degrees of usage. This complicates the instrumentation since it mast measure not merely "how many times," but also "how much" does "how many times" occur. One way of getting this data is to use an oscillograph.

To convert an oscillograph record into useable information, traces are analyzed to determine the frequency-of-occurrence of various magnitudes over the entire route for which the oscillogram was made. Fig. 8 is a plot of frequency-of-occurrence information taken from an oscillograph record.

Having the frequency-of-occurrence curves, a realistic laboratory test is devised by selecting a program of operation that will duplicate or approximate the curve. If conventional testing machines are employed, the usual method of approximating the frequency-of-occurrence curve is to use a step function as diagrammed in Fig. 8.

Recognizing that the order of application of strain, from low-to-high or high-to-low, has an effect on fatigue life, the random distribution of loading which occurs in actual road operation also must be simulated. Generally, this is done by selecting a program of loading that will be repeated many times during the test, and by running the program in both directions as illustrated in Fig. 9.

Transforming hundreds of feet of oscillograph records to frequency-of-occurrence data is a very cumbersome process requiring many man-hours. Enough data usually can be collected in a few hours of road testing to keep a trained analyst busy for days and even weeks.

To overcome the analyzing problem, an electronic selective threshold analyzer was developed. This instrument consists of several counters that total the number of times the load amplitude exceeds certain preselected values. Upon completion of the road test, the numbers on the counters are recorded, from which load versus frequency-of-occurrence fatigue load curves can be plotted directly.

The original selective threshold analyzer has limitations which make it useful for analyzing only low-frequency road strain measurements is a pressing requirement (Fig. 10).

It does not appear feasible, at this time, to establish once and forever the duty cycles for all categories of components. The influence of changing vehicle designs is such that we cannot specify the loads and cycles to which a particular load-carrying member will be subjected. The alternative is to develop a simple and rapid procedure for quickly establishing the load program on a proving ground check run. The relationship between test cycles in the laboratory and required customer service life is then found from the 2.5/1 ratio between proving ground and customer mileage.

Performance Testing

In the area of performance testing, effort in the past few years has been directed toward establishing reproducible measures truly representative of conditions which a large number of customers will experience.

Fuel economy is a major concern of the customer, for obvious reasons. In comparing alternate designs or establishing our competitive position in this respect, we often would prefer to measure only one value of "miles per gallon." This number should reflect the economy which could be expected by the majority of customers, and should reveal the relative merit of different vehicles.

Unfortunately, fuel economy is not a factor, like weight or price, that can be adequately characterized by a single number. No one average value can describe what the majority, or even a small number of customers will achieve. Reliance on such an oversimplified description can lead to disadvantageous product decisions, or conceal areas where

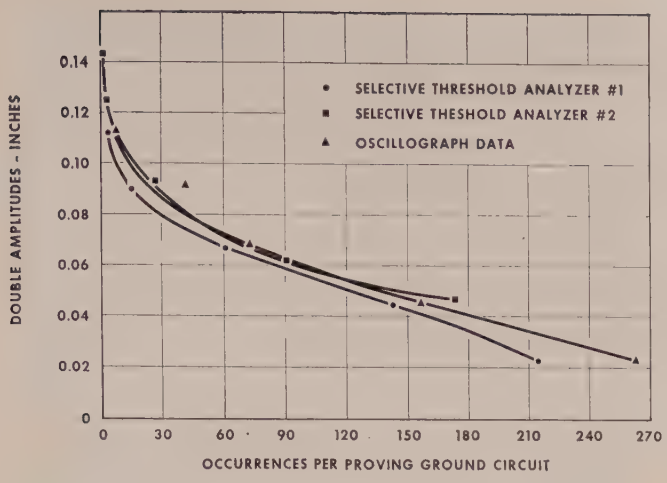


Fig. 10 — Typical vibration characteristic curves, double amplitude versus number of occurrences

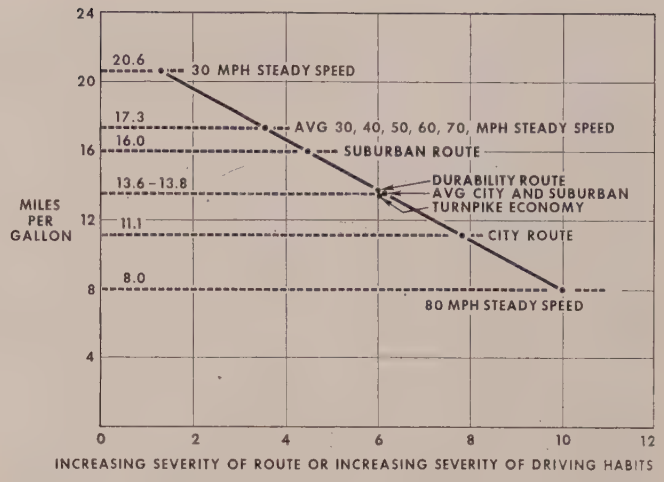


Fig. 11 — Various economy measures, typical car

further attention is needed.

There are five major influences on economy which can affect results to a greater extent than differences between vehicles. These include, in approximate order of importance: speed, changes in speed, route, weather, and the driver. Fig. 11 shows fuel economy of a typical car as measured by some of several different procedures used in recent years. Different test schedules can be devised which can show 2 to 1 differences in economy on a particular vehicle. Likewise, customer usage can be found giving similar extremes in economy. Two other points should be noted, however. First, for any one of these test conditions, there is on the order of only a 1½ to 1 spread between the most economical and the least economical of current domestic automobiles. Usage can cause more difference in economy than design. Second, it is unusal for any one make of vehicle to surpass its critical competition under all of the conditions shown. No single figure, or single average of a group of figures, can show correctly the relative merit of two vehicles, unless we have included in our average all representative types of usage, and weighted them properly.

We must remember that our simulated routes are nothing but assumptions — the assumption that circling downtown streets represents what a large number of drivers do, or that driving from Detroit to Chicago is a typical and valid sample of customer experience. Nevertheless, there is some support for the particular measures in current use which we would like to present.

For many years the automotive industry used — and still uses — what we call a steady speed economy test. This test, run at various constant speeds with stabilized temperatures, on level roads, eliminates the driver influence and is the most reproducible road test.

Unfortunately, the steady speed test yields results much higher than an customer will get. The effect of weight, vehicle inertia, transmission efficiencies during accelerations, accelerator pumps, and warmup also have been ruled out as variables along with the driver. Consequently, this test is supplemented with actual runs in the city, on the highway, or over proving ground durability routes, which give lower and more realistic economies. These tank economies, however, are highly variable due to traf-

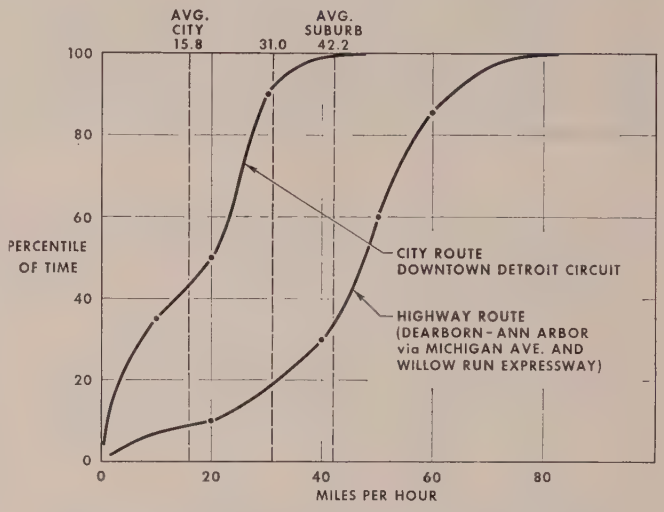


Fig. 12 — Percentile speed distribution, city and suburban routes

fic and road conditions as well as driver habits. Differences between different designs are masked. Comparative tests must be run in convoys, and evaluations from one day to the next are of little value. So these tests have been set up on the proving ground, where there is no traffic interference, and rigorous speed-time, stop, and acceleration schedules can be maintained.

Fig. 12 shows a plot of speed distributions on a public highway course that has been used for customer-type economy checks in downtown Detroit and from Dearborn to Ann Arbor. This route, together with information on starts, stops, and acceleration levels reached, formed the basis for the city and suburban economy procedures we now use. Percentile speed distribution, incidentally, matches very closely similar speed-time records obtained on downtown and suburban areas in Los Angeles, and on other roads and streets in the Detroit vicinity. While considerable departure from this mean curve occurs with daily traffic variations, we feel it is typical of traffic patterns in and around large cities.

Note that use of an average city-suburban route with an overall average speed of 31 mph would suit neither city nor suburban drivers. Ninety per cent of the time city drivers would not reach the average

speed, and over 80% of the time suburban drivers would exceed it. The percentage of time drivers operate at or around the average speed will be small.

Fig. 13 compares the results of a series of suburban economy runs over several 5-mile local stretches of road, with the suburban portion of the simulated proving ground route. Results on the simulated proving ground route are shown as dotted lines. It is evident from this chart that the pattern of starts, stops, accelerations, speeds, and fuel economy on an actual highway run varies widely. These 26 checks were made on a single car, in the space of two days time, and do not reflect total traffic variations. Obviously, if any compartive testing is to be done, it must be done on private roads with full control of the driving program. It is also apparent that the economy obtained on such a proving ground simulated route is close to that obtained on these Southfield and Telegraph checks, as well as on the Expressway and Michigan Avenue which it was designated to reproduce.

In establishing the city and suburban routes in the proving ground, a total length of 8.8 miles was chosen. This is about the median length of trip for the country as a whole. The trend in driving habits has shifted in the past few years, to where

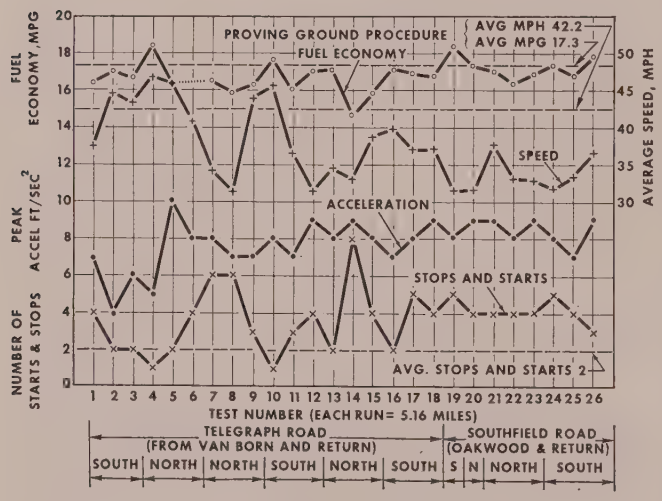


Fig. 13 — Comparison of simulated proving ground suburban economy route and actual highway checks, typical car

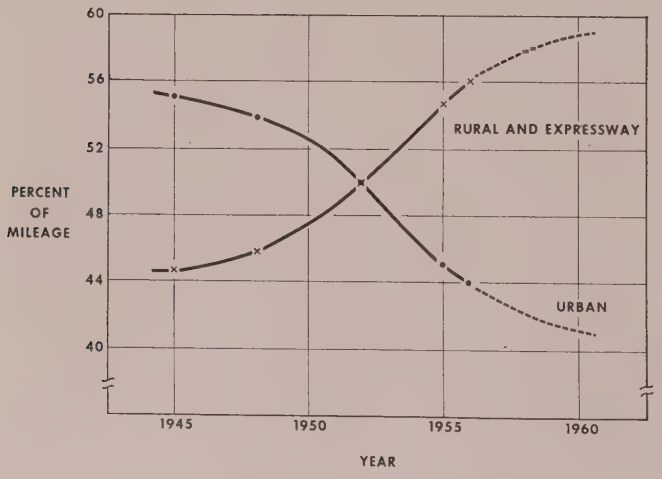


Fig. 14 — Trend in driving habits

the largest part of the car mileage is accumulated on rural roads, and expressways — estimated at about 60% at present (Fig. 14). Our proving ground city and suburban test route was laid out with 60% of the mileage (5.20 miles) on a suburban schedule, to match this current mix of customer usage. Thus, the average of the city and suburban economies obtained should be a properly weighted average.

However, a satisfactory average could be achieved with a low suburban and a high city economy with consequent dissatisfaction for most cusomers. Only if automobiles are acceptable in both areas can we be assured that we are not at a disadvantage with respect to a large number of owners.

In recent years, there has been a trend to higher speed on rural roads and expressways. With many controlled access superhighways already in use, and with funds committed for an interstate highway system similar to the present toll turnpikes to be developed in the next 10 years, this operation will increase.

There are many highway studies which show that sustained high speeds in the 70-mph region are normal to turnpike operation, and that car speeds on these roads are remarkably steady. While the actual mileage accumulated at speeds of 70 and above still is low — only about 1 billion miles annually at present — this is expected to increase appreciably — perhaps to 30 billion miles by 1963. These fast miles will be accumulated mostly on longer trips, where the customer is especially sensitive to his gasoline economy.

Consequently, we feel the city and suburban economy data should be supplemented by a higher speed economy figure — a turnpike economy value — which we feel is satisfactorily simulated by the 70-mph steady speed economy obtained on the proving ground. Fig. 15 shows what these three economy parameters would look like. The heavy car is relatively better, the light car relatively worse in turnpike-type operation.

Methods of measuring and presenting vehicle acceleration performance data have undergone similar evolution in the past two years. Previously, data were obtained to show acceleration rate as a function of vehicle speed. While this had some engineering value, such acceleration curves were not convenient for evaluating comparative performance. Measurements of speed as a function of time also were made; but again, the relative performance of two vehicles was not readily apparent. Consequently, distance travelled from a standing start — in 4 and in 10 sec — have come to be accepted criteria of "getaway." This at least will place cars in their correct relative position, as the customer would see them if he were to engage in a sprint from the stop light.

These two measures give only one view of acceleration as the customer uses it. Passing performance is another vital requirement, and a car with excellent start-up can be very sluggish in passing at highway speeds. The measure we now use to evaluate passing ability is the danger time, or time required to gain certain standard distances over a steady speed car. These exposure times are presented for several initial speeds, usually 30, 40, and 50 mph. Since most people estimate clearances in feet rather than in seconds, the total distance travelled in the exposure time necessary to com-

plete a passing maneuver is also used as a criterion.

There are two other common performance measurements which relate directly to highway situations. One of these is grade climbing ability, which we find we can determine precisely by tractive effort, or drawbar, measurements on level roads. The force applied to a car by a towing dynamometer is directly analagous to the downhill component of the weight of a vehicle operating on a grade. Tire, wind, and various mechanical efficiencies are identical in the two situations. Thus, a drawbar effort versus speed curve with the ordinate expressed in per cent of vehicle weight instead of pounds pull indicates directly the limit of grade capability at each speed, or if you wish, the maximum speed possible on each grade.

The ability to simulate grade effect by application of controlled drawbar loads has been useful in the development of realistic engine cooling procedures for passenger cars. Full-throttle cooling tests on cars are unrealistic with current power-to-weight ratios, automatic transmissions, and improved highways. There are no highway features which require full throttle for more than a few seconds, and the heat capacity of the system will accommodate these transient peak thermal inputs. Our present test procedure for cooling assumes that the continuous heat input is limited to that resulting from negotiating certain standard highway gradients. Minimum highway standards allow 10, 8, and 6% grades for design speeds of 30, 45, and 60 mph, respectively. These grades and speeds, together with 75 mph on level roads, are considered representative operating conditions where stabilized temperatures might be reached. There are such grades in the vicinity of our Arizona Proving Ground — Indio Hill, Townes Pass, Davis Dam Hill — and experience has been that cars which operate satisfactorily in the 100 to 110 deg ambients common in these areas are satisfactory elsewhere.

That outlines the background of our most familiar performance measures. In comparing data from these performance tests, however, and especially in correlating projected vehicle performance with subsequent road tests, it is necessary to consider the effects of weather.

The two main weather influences on performance test results are ambient temperature and barometric pressure. These influences affect the vehicle in two ways; they produce variations in wind and running friction, or power required to move the vehicle; and they produce variations in engine breathing capacity, or maximum power output.

There are well-established methods for evaluating temperature and weather effects on both power requirement of the vehicle and power output of the engine under steady speed conditions.

Examination of results of hundreds of road load power requirement tests at a variety of temperature and pressure conditions show that the vehicle resistance is inversely proportional to absolute temperature, and proportional in part to atmospheric pressure. All of the wind, tire, bearing, and rolling losses appear to be affected in the same way by temperature; only the wind resistance is affected by the air pressure. An equation has been developed which adjusts test data satisfactorily.[1]

Fig. 16 is a comparison of road power data on a 1958 Ford 6 under widely different temperature and pressure conditions, with both sets of data corrected to a common standard. In this case, incidentally, the pressure correction is small, and only about one-fourth as great as that due to temperature.

Tests have also been run on the road many times using driveshaft torquemeters, and results compared with engine dynamometer checks of the same engine. Use of the standard SAE correction formula for engine power have satisfactorily reconciled dynamometer and road test data. Fig. 17 is a typical curve picked at random as an example.

Application of power correction formulas to road performance tests is complicated by two factors. Normally, engine power is not measured on acceleration tests, and the relative effect of weather on both power requirement and power developed varies continuously through a test run as a function of vehicle speed. Nevertheless, a method has been found to apply weather corrections which is theoretically sound and accurate.

Briefly, the technique involves running through a calculation of the particular performance parameter to be corrected, using known vehicle char-

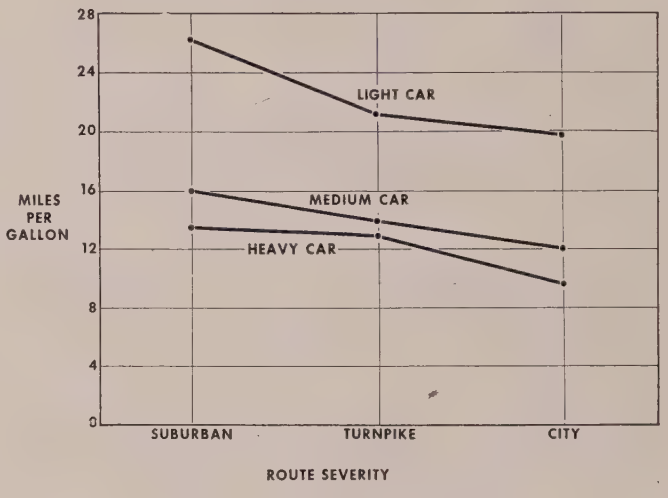


Fig. 15 — Fuel economy comparisons

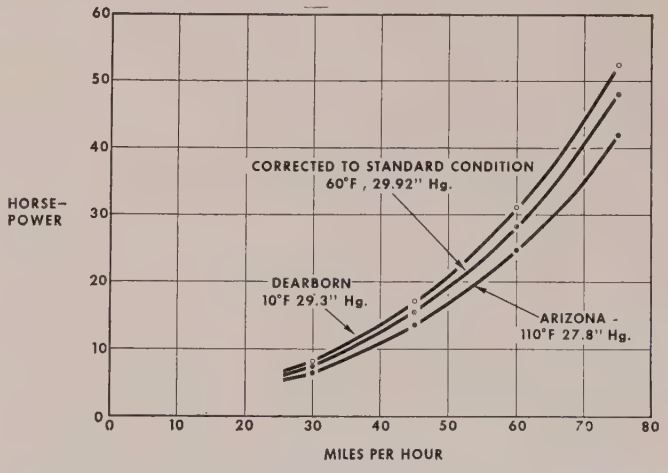


Fig. 16 — Variation of road power requirement with ambient conditions, 1958 Ford 6

[1] Power required, standard conditions, =

$$\frac{\text{Mph}}{375}\left[F_{\text{Rolling}} + F_{\text{Air}}\left(\frac{\text{Bar. standard}}{\text{Bar. observed}}\right)\right] \times \left[\frac{460 + \text{Temp. observed}}{460 + \text{Temp. standard}}\right]$$

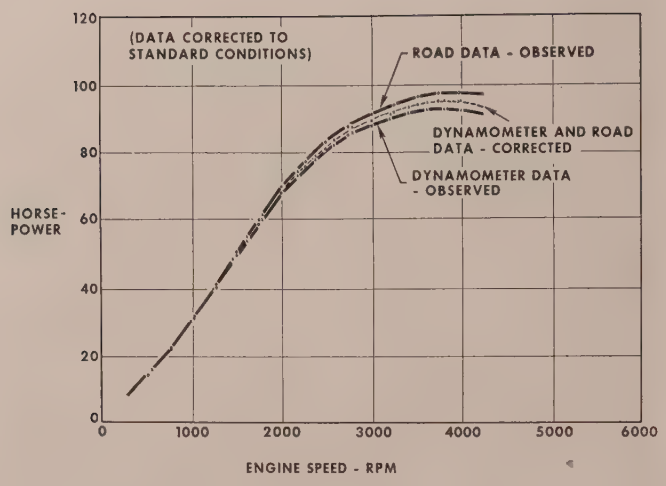


Fig. 17 — Comparison of full-throttle horsepower, typical car, third gear

NUMBER OF CARS TESTED		1	2	3	4	5
1	A	±3.8 %	3.0	2.7	2.5	2.4
	B	±2.7 %	2.3	2.2	2.1	2.1
	C	±1250	1400	1600	1800	2000
2	A	2.7	2.1	1.9	1.8	1.7
	B	1.9	1.7	1.6	1.5	1.5
	C	2450	2800	3200	3550	3950
3	A	2.2	1.8	1.6	1.5	1.4
	B	1.6	1.4	1.3	1.2	1.2
	C	3650	4200	4800	5350	5900
4	A	1.9	1.5	1.4	1.3	1.2
	B	1.3	1.2	1.1	1.1	1.1
	C	4900	5600	6350	7100	7850
5	A	1.7	1.4	1.3	1.2	±1.1 %
	B	1.2	1.1	1.0	1.0	±0.9 %
	C	6100	7000	7950	8850	±9800

NUMBER OF REPEAT TESTS PER CAR

Fig. 18 — Accuracy of test data versus test cost for various sizes of sample, various numbers of tests per sample (A — data accuracy, uncorrected; B — data accuracy, corrected; C — cost of test)

acteristics and observed test weather conditions. The same calculation is then repeated, changing only the weather conditions, to the desired standard. Comparison of the two calculations yields a difference which is applied to the test result as an additive correction. Since vehicles of the same performance level are affected to about the same extent, it has been possible to develop correction tables for each performance criterion for various temperatures and pressures.

At the present time, fuel economy correction techniques need refinement. While the major effect appears to be the effect on road load power requirement, there are significant secondary effects of temperature, pressure, and humidity on carburetion and air/fuel ratios.

We have not yet found a sound theoretical basis for correcting engine cooling data to a standard condition. When tests are conducted to determine the ambient temperature required to produce boiling, obviously the engine power output and heat rejection should be adjusted to whatever the value would actually be at this air-to-boil temperature.

If cooling tests are run at some lower temperature, the engine output will be higher, heat rejection higher, and the calculated air-to-boil figure will be too low. During wind tunnel tests we have found an empirical relationship which says that for every increase of 1 hp in engine output, there will be a decrease of 1 deg in the air-to-boil figure. This relationship appears to hold approximately for all the part-throttle loads used on passenger cars, and for full-throttle tests as well. It is a useful relationship in evaluating road cooling tests, and reconciling differences between cooling tests run under different weather conditions.

Accuracy of Test Data

Even with corrections or allowances for weather variations, there are limitations to test data which must be recognized, especially when comparing different makes or types of vehicles. Automobiles — even those that are nominally identical — are very much like peas in a pod: all slightly different. To establish characteristics of the species, it is better to measure them all, or as many as we can, with even a crude caliper, rather than to take a particular pea and examine it with the most precise tool. To have confidence in the reliability of our test data, we must consider both the variability between individual samples as well as the variability of our observations of those samples.

We might make some fine distinctions here between repeatability of measurements, accuracy, and possible spread of test data.

We can get repeatable or reproducible results using a 2-oz measuring glass in the mistaken belief that it holds 1 oz, but our accuracy will be poor. Measuring accuracy depends on the precision of our test instrumentation, and on how carefully it can be calibrated against primary standards. We are limited on road test performance measurements by the resolution and reproducibility of the best apparatus we can devise to only about 1% on measurements of speed, time, volume, distance, mass, and temperature. Within what we feel to be practical limits of time and cost we are not likely to do much better. Fortunately, these errors are not additive; we can also estimate their cumulative effect from repeat measurements. In comparing vehicles, the repeatability of our observations is of more concern than their absolute accuracy.

Spread or scatter of test data, like the size of peas in a pod, depends not only on measuring errors, but also on the adequacy of the test sample. The more vehicles used in a sample and the more observations we make on each vehicle, the more different answers we will get, and the wider will be our spread of data. The wider the band spread, the greater will be the probability that the next sample tested will fall within this band. It also follows that the wider the spread, the closer the average of these data will be to the true mean which would be established by accurate measurement of the total population.

We express accuracy of our test data, therefore, as a per cent range within which there is a certain probability that the true value will be found.

Fig. 18 is a chart showing the accuracy we can expect on one of our performance test procedures with various numbers of cars in the sample, and with various numbers of observations on each sample, at an 80% confidence level. Thus, if we want results that will be accurate to better than 2% four times out of five, we will have to test at least

two samples. If we want better than 1% accuracy, we must make five tests on each of five samples. If we want to work with better odds — say 19 times out of 20, or a 95% confidence level — all the accuracy percentages shown will be increased to about 1½ times the values in this particular table.

Cost data such as is also given in Fig. 18 helps us decide how much accuracy we can afford. Doubling accuracy roughly quadruples cost. It is also apparent that procuring and preparing additional samples is more costly than increasing the number of tests performed. There are other considerations that are sometimes dominant — for example, available manpower or time allotted. Inclusion of all these factors in a table allows us to choose a test program which will give the best results compatible with practical limitations and economic justification.

Accuracies are shown in Fig. 18 both with and without weather correction applied. The utility of correction is perhaps greater than the differences between the figures would suggest. The series of tests from which these accuracies were derived were performed before we had much faith in our correction techniques. Consequently, the data were all accumulated at one proving ground area with a relatively narrow barometric pressure variation, and only on days when ambient temperatures were within a 45–75 deg range. We now feel we can allow substantially more latitude in weather, so that many more days will be suitable for performance testing, and we can be less concerned over the altitude and pressure differences between our three proving ground locations.

The basic information used to develop accuracy tables for each of our vehicle performance parameters (acceleration, passing ability, city and highway fuel economy) was derived from a fleet of 45 cars of our own and competitive makes with up to 12 test observations per car. In most cases, there were three vehicles of each type, and four tests on each of these three vehicles. From these data we could establish a total variation including both sample differences and observational error. From the repeat tests on particular samples, we established variations attributable to observational error only. The difference between total variation and observational, or test variations gave a variation due to differences between samples. Through standard statistical formulas,[2] the car variations and test variations were combined to find a total standard deviation for other combinations of sample size and number of observations. It is interesting to note that the variability between different samples of the same type car was surprisingly uniform for all the different makes and performance levels of cars tested.

Conclusion

We are aware that this paper has not directly answered the question raised in the title — "How Good is Testing?" It has been, rather, a progress report, outlining the state of the art, and indicating areas which will need further, or continuing, attention.

Comparison of warranty data, dealer repair orders, or other sources of information on customer experience, with laboratory and proving ground tests of production vehicles will be a continuing necessity, if we are to be assured that these test operations accurately reflect customer usage.

As customer needs and desires evolve, the measures we apply to describe performance, economy — how well the vehicle fulfills its function — must change accordingly.

As product improvements are sought in smaller and subtler ways, we must further sharpen our tools — improve precision of measurement, assess the influence of driver and weather variables with greater accuracy, and employ statistical techniques — to confirm the differences.

Perhaps the only answer possible to our question is that testing will never be quite as good as we wish.

---

[2]

$$\sigma_{Total}^2 = \frac{\sigma_{Car}^2 + \frac{\sigma_{Test}^2}{N}}{K}$$

for N = Number of tests on each of K number of cars

where:

σ_{Total} = Standard deviation including sample and test variations

σ_{Car} = Standard deviation of multiple samples

σ_{Test} = Standard deviation of multiple observations

DISCUSSION

Testing Must Be Compared With After-Warranty Experience

— A. G. Loofbourrow
Chrysler Corp.

IN EVALUATING proving ground test results against actual service experience, Mr. McConnell uses field information reported during the 4000 mile warranty period. However, a complete answer to the question, How Good is Testing?, must also consider customer experience beyond the 4000 mile warranty period.

The problem of correlating endurance test results with after-warranty customer operation is one that the industry must accept in the interest both of its customers and of its own reputation for producing durable vehicles that continue to perform well through the years for the second, third, and even fourth owner. If we do not, if we merely set out to anticipate the type of failures that occur in the first 4000 miles of customer operation, we are open to the charge that our only interest is in protecting our own warranty money, not in minimizing the after-warranty failures for which the customer must pay.

Two major obstacles face us from the start. The first of these is the task of simply finding out what the after-warranty field service picture may be. This is no easy matter. Warranty information comes in quickly, since our dealers do not delay overly much in getting in their claims. But there is no similarly direct channel for finding out what is going on with cars once the warranty period is up. The customer brings the car in for service, pays the tab, and that is that. He may not even bring the car to a franchised dealership, for that matter. There are several sources of clues to this riddle. First, we can watch the sales of service parts to catch any unusually fast-moving items. Large fleets are another good source of after-warranty service information, since mileage on these cars will be accumulated at a much faster rate than on the average car. Perhaps the most obvious source of this information is our dealers. We do get an indication of some problems through dealer council meetings and from our field force. However, neither channel is quite as revealing as a good down-to-earth talk between an engineer and an alert service manager. At Chrysler we are now doing this on a full-time basis.

The second of our obstacles arises when we try to work out a procedure for alerting ourselves to potential after-warranty failures through laboratory and Proving Ground

testing. As Mr. McConnell has recognized, we are faced with the job of trying to accelerate things which don't like to be accelerated: Aging, corrosion, weathering, and soaking. It is the responsibility, however, to go beyond the accepted laboratory tests for these factors and attempt to obtain advance information based on customer-type operation.

Field testing, in which new components are installed in taxicabs or other fleet cars, is a valuable source of this information. In some cases, we have installed experimental parts on our employees' cars, thus not only gaining customer-type miles, but the benefit of an accurate history of use by trained observers. We are getting good use out of a special corrosion cycle which we recently established at our Engineering Proving Grounds, in Chelsea, Mich. We have worked out special soaking schedules, and other means of duplicating the service environment of the high-mileage customer's car.

We do not consider that we yet know all there is to be known about this area of endurance testing, but we have made a start. It is certainly an area in which continued work can result in direct benefit both to the industry and it customers.

Before leaving the subject, I think that lest we create the impression that the industry is designing and building cars that continually drop broken parts along the highways, we had better define the term failure as used in connection with endurance testing. To the test engineer, a part is considered to have failed when it no longer functions satisfactorily as the designer intended. The part might begin to get noisy, it might develop a leak, or it might operate less efficiency than expected. A failure does not necesarily mean the part is no longer intact, nor that the car itself has been immobilized.

As a matter of fact, in some instances the line between satisfactory performance and a failure is not entirely clear. For example, if noise develops in a component we may still ask the question: Is this noise within the so-called "commercial" area, or is it bad enough to require redesign? By keeping our standards of what constitutes "commercial" performance high, and by an intensive program of endurance testing closely tied to actual customer experience both during and beyond the warranty period, the industry can continue to enhance its reputation for providing the best automotive transportation values in the world.

Proving Ground Data Helps Predict Customer Service

— Louis C. Lundstrom
General Motors Corp.

THE FORD engineers have done a commendable job in constructing their Michigan proving ground roads to a prearranged severity. Several of us at the General Motors Proving Ground have had a somewhat easier job because we inherited a road system that was started in September, 1924, and developed through the years to provide a degree of severity that would predict customer experience in a reasonable period of time and number of miles.

While it has been desirable to maintain as consistent a test program as possible over the years, the automobile has improved remarkably, and the customer's driving habits have changed sufficiently to make it necessary to increase the severity of the schedules. This is, in effect, proof that over the years the product has been increasing in service and dependability. Top speeds for our endurance schedule have been gradually increased. The Belgian Block Road was added to the schedule several years ago. Additional brake stops have been added. Backup or reverse operation, both on level roads and on the hills, has added to the evaluation of the transmission. Corrosive salt sprays are now a part of the standard schedule. Additional operating procedures, such as starting and stopping the engine, opening and closing doors and windows, playing the radio, and even blowing the horn, have become a part of the test schedule.

As an interesting sidelight, confidential chart rooms and displays of durability parts were used in the late 1920's in place of detailed engineering reports. Early parts classification systems were then being developed. It is also interesting to note that a large backlog of test information was accumulated at our Proving Ground before Service Section records became available.

The comparison of service reports and endurance testing by Ford is interesting and well presented. A few general comments may be of interest on this subject. Perhaps the most serious problem with the service reports is that the records cover only a limited number of miles of operation. They emphasize the manufacturing quality and not the long-term life built into the modern automobile. Our Proving Ground endurance testing, in three months, produces wear or part failures that may not be encountered by the first or even by the second owner of the car over several years of normal customer operation. The 25,000- and 35,000-mile tests normally used represent approximately 50,000–70,000 miles of operation by the average owner. Field engineers play an important role in collecting customer experience early in the model year.

Mr. McConnell points out that, "Only a limited sample of production vehicles normally are tested over the proving ground roads, to get a direct comparison of customer versus proving ground failures on identical vehicles." I do not know the number of samples involved, but, again, the many years of testing on our Proving Ground have provided a large file of information to support the validity of our endurance or durability test.

Our own testing does not confirm Mr. McConnell's statement that "About as many failures occur up to 20,000 miles as occur between 20,000 and 40,000 miles." We expect to replace items such as brake linings and tires, but many other parts are expected to run the life of the vehicle. We have many times extended the standard tests beyond the 25,000-mile period to obtain failures not found at lower mileages. Because of lack of time, we usually keep the standard mileage and rely on inspection of parts for wear and impending failure, rather than attempting to run every part to complete failure. As stated before, additional failures can be obtained at the same mileage by increasing the severity of the test conditions.

Mr. McConnell only briefly mentions the use of severe schedule to accelerate the testing of new designs. This is often done to evaluate a new design in comparison with current production, or to determine the differences between two new designs. Such schedules may require additional Belgian Block operation, additional exposure in the Salt Trough, additional trips through the Mud Bath, additional high-speed stops, or higher speed operation. Here, the engineer's skill in interpreting results becomes increasingly important. Oftentimes, his conclusions and recommendations to management will determine the serviceability of the vehicle in the hands of the customer.

Service is performed on endurance cars in accordance with owner's manual instructions. Parts that fail because of owner neglect cannot be charged against design.

Checking the performance of the endurance cars at intervals during the test determines how well the car retains its original condition.

The durability schedule generally provides more of everything than the average customer experience: More dirt and rough roads, more hills, more high-speed operation, more turns, more stops, more transmission shifts, more reverse operation, more night operation.

We certainly concur with Mr. McConnell that there is no one number to define gasoline economy adequately. In the present economy race, the public will be just as confused as they were during the horsepower race. SAE has for many years had a definition for horsepower, but so far as I know, they have never attempted to define economy. Ford's use of the level road economy, city economy, suburban economy, and turnpike economy, is undoubtedly satisfactory for their engineering and management's use.

We have developed level road, city, and highway schedules that provide our engineers and management with similar information, but the schedules and results are not expected to duplicate those of any other manufacturer.

In general, I believe that the development engineers, with the help of proving ground data, are doing a good job of predicting customer service. If they were not, the warranty and policy costs would soon become excessive. The variations in driver habits and driving conditions are so extreme, however, that every condition cannot be duplicated or evaluated. The company and dealer organizations can in many cases assist the buyer in obtaining the particular vehicle and options that will give him the best overall satisfaction.

Prototype and Production Models Both Important in Durability Testing

— John F. Adamson
American Motors Corp.

IN READING this paper, I cannot help but notice some of the differences that exist within the industry not only in relation to test procedures, but also in the results obtained. Conceivably, this is more noticeable to someone like myself, who is used to dealing with proving ground and laboratory findings that are peculiar to the unitized body type of construction. Of course, the idea that various differences in car construction require some modification in testing conditions in no way alters the fact that all members of the industry are striving for the same goal. Obviously, the primary aim of everybody is to give the public an automobile that will prove to be a durable and satisfactory vehicle.

With reference to these differences that I mentioned, we must bear in mind that in a car with the unitized body type of construction, a number of the running gear or chassis components are tied directly into the body structure. As a consequence, it is of the utmost importance that these proving ground procedures include an adequate amount of the kind of durability testing that is sure to provide the necessary straining and working of these structures. It is through the type of research that Mr. McConnell mentioned that we have determined just what kind of durability and deflection testing is necessary to represent a normally severe type of usage for this type of construction.

Training personnel to examine cars so tested is also an important part of "How Good Is Testing," as a failed part means nothing unless some one can detect it and an intelligent report is given to the interested engineers.

One of the major thoughts brought out in the paper, which I think, is of the utmost importance, is this idea of thoroughly testing production cars in addition to prototype models. The basic design of a part can be put through its paces on a pre-production prototype, but there are times when the minute differences required in production pieces can seriously reduce the life of the part in question. Unless these differences are taken into account, any test program is somewhat worthless, as it is not built around the structures and assemblies that the ultimate customer is going to be riding around on.

This can best be illustrated by noting that there can be a considerable difference in the strength and durability of a prototype engine or suspension crossmember when compared with its production counterpart. The former would possibly be made over wood or plastic hammer forms and fabricated by arc welding, while the latter would be die formed and spot-welded together. With regard to this, I think that we can all visualize the differences in stresses due to the two methods of metal forming. In addition, I think that we would also agree that there can easily be a difference in weld strength between components arced together and the same assemblies utilizing production spot-welding techniques.

Still another point brought out is the problem confronting automotive engineers in determining just which testing failures represent what we can expect in the field. I thought that the presentation very graphically showed that this is a subject on which a great number of man hours have been spent in correlating data obtained at the proving grounds and in the laboratory. I am very pleased to see that some of the great amount of work being done behind the scenes is now being made public knowledge.

There is one point in Mr. McConnell's discussion, however, with which I cannot agree. While discussing endurance testing in the laboratory, he stated "There are many items that are of the 'on-off' or 'open-shut' variety which require only knowledge of 'How Many Times' to set up laboratory tests. Most electrical switches or controls, doors, deck lids, and window and hood operating mechanisms, fall into this category."

Actually, I don't think that electrical switches can be tested quite so simply. Fairly recently, we ran into field problems with the starter switch on our nonautomatic transmission equipped cars. This switch was of the ignition key type, and although we were getting reports of field failures, we could not duplicate these failures in the laboratory. We finally discovered that although these switches would successfully cycle many more times than required on a simple 'on-off' test, the driver was not actuating the switch in the same manner. Instead of rotating the ignition key to the full 'starter-on' position, there were times when he would either rotate the key only far enough to start the cranking motor, or he would go to the full on position and then slowly back off while current was still going through the switch. This, of course, caused internal arcing and the early failure of the unit. It was only after this discovery, and the subsequent laboratory test revisions, that we were able to say that we were testing to field conditions and actually determining the normal life of the switch in question.

In conclusion, the subject "How Good Is Testing" is one that certainly occupies a prominent spot in the automotive industry. It is a part of this business that assures our customers of a satisfactory and well-proved vehicle.

Author's Closure To Discussion

MR. LOOFBOURROW mentions a point that is extremely important — the fact that the customer is highly interested in what happens to his purchase well beyond the warranty period. This is also of great concern to the manufacturer, if only because it has a large influence on resale value, and hence on sales. The sophisticated new car buyer is well aware that his trade-in allowance several years hence may largely determine what his car has cost him to own. We would, therefore, like to clarify an impression we may have given that we are primarily interested in reducing warranty cost.

We used warranty information because it gave us the largest mass of accurate data on customer difficulties, so that we could determine distributions of troubles, not only in the 35 UPC groupings, but also in the numerous subgroups within these main categories. Matching the type, frequency, and distribution of failures experienced by all our customers in the short warranty mileage assured us that our proving ground operation was realistic. Thus, we are confident that our 20,000- and 40,000-mile proving ground schedules are equivalent to 50,000 and 100,000 miles of customer service.

Every manufacturer, of course, supplements his proving ground testing to fill the gaps. These supplements include the accelerated corrosion tests, used car lot inspections, service part sales, fleet owner records, dealer councils, garage repair order analyses, and customer surveys. We would like to be sure that subsequent owners, as well as the original buyer, are spared both the defects that might affect the safety and utility of his transportation, and the far more frequent minor annoyances and disappointments that we also classify as "failures."

THE LOW-POWER SPACE NUCLEAR ROCKET conceived by NASA engineers is described in this paper. It is compared with the chemical rocket and the nuclear turboelectric ion propulsion system.

In developing the concept for this low-power rocket, NASA engineers concentrated on attaining low weight and high hydrogen temperature, and on solving problems concerned with automatic control and operation of high-temperature reactors.

It was presumed that the NASA 1.5 million-lb thrust engine would be available, and could place 25,000 lb in orbit, at the time the nuclear rocket is ready for test.

As experience is gained reactors of higher power can be developed. These can, perhaps, be used as second stages of larger chemical boosters. Finally, high-power, high-temperature rockets for booster application can be undertaken.

In addition to describing the low-power nuclear rocket, the authors discuss the various types of orbits possible for earth-Mars trips and the performance possibilities of nuclear rockets.*

Nuclear for

Frank E. Rom and Paul G. Johnson
National Aeronautics and Space Administration

EVER SINCE the first atom bomb was exploded there has been a great deal of speculation about the use of nuclear energy for rocket propulsion. Because of the high energy content of uranium (about two million times as much as chemical propellants) nuclear energy appears to hold the key to efficient space exploration.

Only in the last three or four years have serious experimental efforts been undertaken to physically demonstrate the feasibility of nuclear rocket propulsion. Los Alamos Scientific Laboratory was the first to tackle the problem of constructing an experimental reactor that could be considered as a nuclear rocket type[1]. It is scheduled to be tested in the near future. Other organizations have been doing experimental work connected with some of the particular problems of nuclear rocket propulsion. Now that our country is placing a great emphasis on space exploration, the interest and effort expended on nuclear rocket research has increased greatly.

Space Missions

Some of the kinds of space missions in which there is interest are illustrated in Figs. 1 and 2. The missions are divided into two classes: those missions in the vicinity of the earth and the moon are called near-earth missions and those missions beyond the earth and the moon are called interplanetary missions. The near-earth missions are illustrated in Fig. 1. The placing of payloads into low-altitude orbits from the surface of the earth is the first listed. Such missions are typified by the Explorer, Vanguard, and Sputnik series of satellites. They require a rocket system to accelerate the payload through a velocity change of about 5 mile/sec. Another mission of interest is the establishment of payloads in a 22,300-mile orbit called the "stationary orbit." It derives its name from the fact that the period of revolution about the earth is one day, so that the payload appears to be stationary with respect to the surface of the earth. This orbit requires a velocity increment of about 7 mile/sec. Still another mission of interest is the lunar probe starting from the earth's surface in which a payload of scientific instruments is placed in the vicinity of the moon. This mission also requires about 7 mile/sec velocity increment.

The stationary orbit and lunar probe missions can also be started from an established earth orbit. The velocity increment for the low-altitude orbit is obtained by the use of one propulsion system called a booster, while the remaining increment is delivered by a second propulsion system called a space engine. For example, the stationary orbit and the lunar probe require only about 2.4 and 2.1 mile/sec additional velocity increments. In other words, if the space vehicle can be placed into orbit, the task of completing the mission is greatly simplified. Another important advantage should be pointed out. Once in orbit no thrust is required to maintain altitude, which means that a minute thrust, only a small fraction of the vehicle weight, can be

* Paper presented at SAE National Aeronautic Meeting, New York, April 3, 1959.

[1] "Los Alamos' Project Rover," by Raemer E. Schreiber. *Nucleonics*, Vol. 16, July 1958, p. 70–72.

Rockets

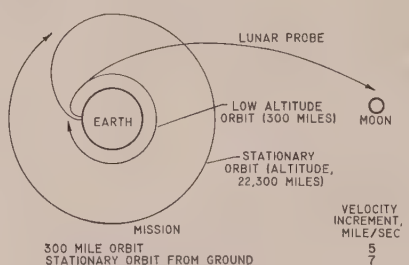
Interplanetary Propulsion

used to produce the velocity changes required for space missions. This flexibility broadens the number of propulsion-system types which can be used.

The second class of space missions is that in which the vehicle explores space beyond the region of the earth and the moon but still within the solar system. These interplanetary missions are illustrated in Fig. 2. The first example shown is the minimum-energy Mars probe in which the space vehicle is accelerated away from a 300-mile earth orbit so that it coasts on an elliptic path tangent to both earth and Mars orbits under the gravitational influence of the sun. The time for this trip is 259 days and requires the least possible velocity change (about 2.2 mile/sec). The transfer time between earth and Mars can be shortened at the expense of higher velocity increments. For the 120- and 80-day trips illustrated the velocity increments are increased to 2.8 and 4.3 mile/sec, respectively. To send a probe anywhere in the solar system requires a velocity increment of 5.4 mile/sec.

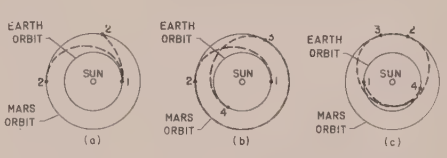
The minimum-energy round trip to Mars uses the 259-day coast as in the probe mission. At Mars the space vehicle is decelerated into a Martian orbit. The vehicle must then wait 455 days until the positions of Mars and earth are favorable for a return trip of 259 days. The total trip time is 973 days and requires a total velocity increment of 6.9 mile/sec. It is possible to find paths which are tangent to neither Mars' nor earth's orbit about the sun. An interesting example of this trajectory type is called a swing-around path. The spacecraft leaves earth's orbit on an elliptical path about the sun and arrives at Mars vicinity after a 226-day coast period. The spacecraft is deflected by the effect of Mars' gravitational field and coasts back to earth on a 282-day path. No propulsion is required at Mars other than for navigational corrections. The spacecraft will be within at least 1,000,000 miles of Mars for a period of about 6 days. The closest approach during this time is 1000 miles. This mission requires a total velocity increment of 9 mile/sec and takes a total time of 508 days.

For missions where it is desired to remain in orbit about Mars for a shorter period of time than 455 days dictated by the minimum-energy trip, excess-energy trajectories must be used. For example, a 270-day path which leaves earth's orbit at an angle, but is



MISSION	VELOCITY INCREMENT, MILE/SEC
300 MILE ORBIT	5
STATIONARY ORBIT FROM GROUND	7
LUNAR PROBE FROM GROUND	7
STATIONARY ORBIT FROM 300 MILE ORBIT	2.4
LUNAR PROBE FROM 300 MILE ORBIT	2.1

Fig. 1 — Near-earth space missions



MISSION	TRIP TIME (DAYS)	VELOCITY INCREMENT (MILE/SEC)
MINIMUM ENERGY MARS PROBE (a)	259	2.2
EXCESS ENERGY MARS PROBE	120	2.8
EXCESS ENERGY MARS PROBE (a)	80	4.3
ESCAPE FROM SOLAR SYSTEM	---	5.4
MINIMUM ENERGY MARS ROUND TRIP (b) (455 DAYS AT MARS)	973	6.9
EXCESS ENERGY MARS ROUND TRIP (SWING-AROUND)	508	9.0
EXCESS ENERGY MARS ROUND TRIP (c) (30 DAYS AT MARS)	570	14.0

Fig. 2 — Interplanetary missions

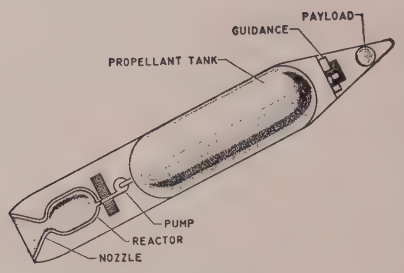


Fig. 3 — Nuclear rocket

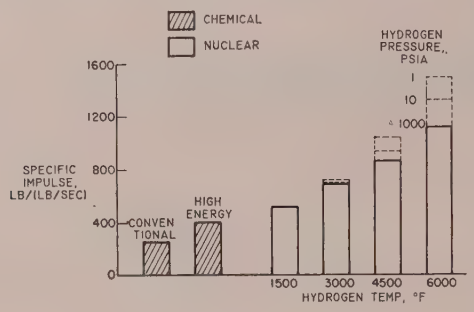


Fig. 4 — Specific impulse

tangent to Mars orbit as shown, permits a wait time of only 30 days. The return trip is made on a symmetric 270-day trajectory. This trajectory is one of a family which has earth and Mars in opposition while the spacecraft is circling Mars. The distance for transmission of data between Mars and earth is close to minimum for opposition. The total velocity increment required for this mission which takes a total of 570 days is 14 mile/sec.

It can be seen that velocity increments required for interplanetary transportation are up to three times as great as that required to place a vehicle into an earth orbit. Chemical propulsion systems are already straining to accomplish even the satellite mission. (The payloads are only about 1–2% of the take-off weight.) Thus, the relatively low energy content of chemical propellants seriously limits their capabilities. Enormous amounts of chemical fuel are necessary to provide enough energy for space flights with useful payloads.

Nuclear Rocket Propulsion

How can nuclear energy, with energy content about two million times that of chemical propellants, be utilized for rocket propulsion? The simplest way is to use it to heat the ideal rocket propellant and eject the propellant through a nozzle. The best rocket propellant is the one with the lowest molecular weight, since it can be ejected through a nozzle at the highest velocity for a given area ratio and temperature. Hydrogen is the obvious propellant to use. Fig. 3 is a schematic drawing of a nuclear rocket. Hydrogen in liquid form is stored in a tank. A pump pressurizes the hydrogen and pumps it through a nuclear reactor, where it picks up heat liberated in a chain reacting mass of uranium-235. The hot hydrogen is then ejected through a nozzle to produce thrust.

Specific Impulse

The performance of nuclear rockets compared to chemical rockets in terms of specific impulse is presented in Fig. 4. Specific impulse, defined as the pounds of thrust produced for each pound per second of flow through the nozzle, is shown for chemical and nuclear systems. Conventional chemical propellants such as gasoline-oxygen produce about 250 lb of thrust for each lb/sec of flow. The best chemical propellants are expected to produce specific impulse values of about 400 lb/(lb/sec). The remaining four bars indicate the specific impulse attainable by ejecting hydrogen heated to various temperatures through a nozzle. At 1500 F a specific impulse of about 515 can be expected. As the temperature is increased to 3000, 4500, and 6000 F, the specific impulse increases to values of 690, 880, and 1160, respectively. The highest temperature shown is about the best that could be hoped for since it is only about 1000 F below the melting point of tantalum and hafnium carbides, the most refractory materials known.

The specific impulse values indicated are for a reactor outlet pressure of 1000 psia. If the pressure is reduced sufficiently, the hydrogen will become dissociated. This will tend to increase the specific impulse for two reasons. One is that the average molecular weight of the gas is reduced, and the other is that recombination of some of the dissociated atoms during the expansion in the nozzle adds energy to the stream, thus tending to increase its temperature. The specific impulse is increased from 1160 to 1330 lb/(lb/sec) at 6000 F if the pressure is reduced from 1000 to 10 psia. At 1 psia the specific impulse is 1500. At 4500 F the specific impulse is increased from 880 to 940 to 1040 lb/(lb/sec) as pressure is reduced from 1000 to 10 to 1 psia, respectively. At 1500 F and 3000 F not much dissociation occurs so there are no large gains to be observed.

From Fig. 4 it is apparent that values of specific impulse much greater than those attainable with chemical systems are possible with nuclear rockets. Even at a temperature as low as 1500 F some superiority is indicated. Operation at low pressures to obtain high specific impulse pays off for missions with large velocity change requirements where the propellant weight represents a very large fraction of the total vehicle weight. This is done at the expense of engine weight, however, since the engine must be larger due to the greater volume flow rate of hydrogen at reduced pressures. Inasmuch as the engine weight in such missions is a small weight fraction, the penalty may not be as important as the gains resulting from the increase in specific impulse.

Applications

Nuclear rocket applications can be divided into two categories: (1) those missions which start from the earth's surface and require thrusts greater than the vehicle take-off weight, and (2) those missions which start from a satellite orbit about the earth and require thrusts only a small fraction of the starting weight.

Power Requirements — Fig. 5 illustrates the power requirements for nuclear rocket boosters for a large range of payloads to be placed into a 300-mile earth orbit. The reactor power is given in megawatts (1 megawatt is equal to about 1300 hp). The payload is given in thousands of pounds. A range of hydrogen temperatures is shown. The large range of payloads is chosen since the higher values approximate the weight requirements for future manned interplanetary flights. For a payload of 100,000 lb, the reactor

must operate at a thermal energy output rate of 10,000–30,000 mw depending on the operating temperature. The higher value is equivalent to 40,000,000 hp. The reactor power requirements go up to 100,000 mw for a 500,000-lb payload. Even for payloads as low as 25,000 lb, the power is about 5000 mw. The use of nuclear energy for leaving the earth's surface requires reactors operating at power levels several orders of magnitude beyond current practice.

Fig. 6 illustrates the power requirements for nuclear rocket space vehicles which start from earth satellite orbits. The reactor power in megawatts is plotted as a function of initial acceleration in g's for initial weights in orbit of 25,000, 100,000 and 500,000 lb. Each curve is shown as a band indicating a spread caused by operating at various temperature levels ranging from 1500 to 6000 F. The largest power requirement is 1000 mw for the 500,000-lb spacecraft with an initial acceleration of 0.1 g. The initial weight of 500,000 lb corresponds to weight estimates required for manned interplanetary voyages. At accelerations of 10^{-2} g, which are adequate for space missions, the power requirement is reduced to about 100 mw for the 500,000-lb spacecraft. At 100,000 lb, the power required is 15–20 mw, while at 25,000 lb only 4 or 5 mw are required. At the lower powers, however, the powerplant must operate for longer periods of time to achieve the desired velocity change. The powerplant must, therefore, be designed with greater durability in mind. Nevertheless, the point to be made is that high reactor powers are not necessary for space propulsion if the vehicle can start from an earth orbit.

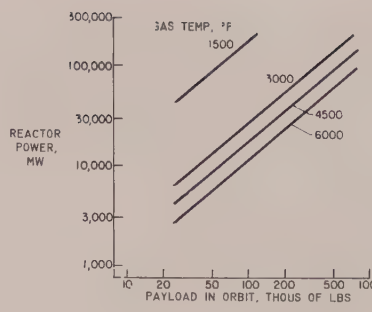


Fig. 5 — Power requirements for booster nuclear rockets

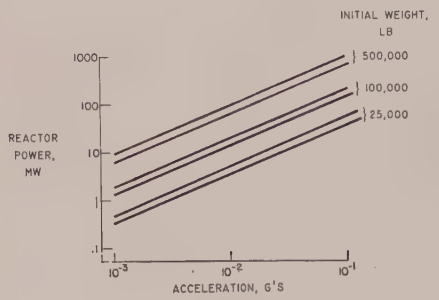


Fig. 6 — Power requirements for space nuclear rockets

Flight Time Consideration — Fig. 6 would indicate that any power at all would be satisfactory. For example, if the curves are extrapolated to an acceleration of 10^{-5} g, it would take only 0.1 mw to operate the 500,000-lb space vehicle. Another factor must be taken into account, however, and that is time. If the acceleration approaches zero, then the time to achieve a finite velocity increment must approach infinity. Fig. 7 shows how the trip time for various interplanetary missions is affected by initial acceleration. The missions shown are the minimum-energy Mars round trip, the excess-energy 570-day Mars round trip, the 508-day swing-around Mars trip, the minimum-energy Mars probe, and the 80-day excess-energy Mars probe. At accelerations less than about 10^{-3} or $10^{-4}g$, the total trip times start increasing rapidly. At 10^{-4} g, for example, it would take about 70 days, or nearly 400 revolutions about the earth, to achieve the velocity required for a minimum energy Mars trip. Above an acceleration of 10^{-3} g, the total trip time approaches the coast time and very little time can be saved by accelerating at higher rates. This is because the acceleration or propulsion time is negligible compared to the coast time.

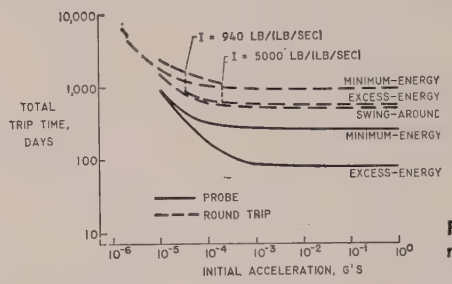


Fig. 7 — Trip times for nuclear rockets (4500 F temperature)

In the case of the minimum- and excess-energy round trip missions, discontinuities are observed in total trip time. For the minimum-energy round trip, 455 days must elapse between the time of arrival and departure from Mars. Out of this time must come the time necessary both to decelerate into an orbit about Mars, and to reaccelerate out of the orbit. If the acceleration is so low that it would take longer than 455 days to decelerate and accelerate again, the return journey would have to wait until earth and Mars are again in a favorable position. This would add 780 days to the trip. This jump in time occurs for an initial acceleration of about 2×10^{-6} g. The jump does not add a significant amount of time to the trip since it is already 4500 days.

In the case of an excess-energy round trip which leaves tangent to earth's orbit and requires a 30-day wait at Mars, the minimum permissible acceleration is increased to about 4×10^{-5} g for the nuclear rocket operating at 4500 F with a specific impulse of 940. The trip time will show a sudden jump of 780 days for accelerations less than this minimum. If the specific impulse is greater than 5000 lb/(lb/sec), such as would be the case for nuclear electric propulsion systems, the jump occurs at accelerations greater than 2×10^{-4} g. The nuclear electric spacecraft must therefore have accelerations greater than 2×10^{-4} g to perform this mission which has a 30-day waiting time at Mars.

Mass Ratio Requirements — Still another factor must be considered before a choice of initial acceleration can be made. In low acceleration flights, the flight path consists of a gradually increasing spiral about the earth. The lower the acceleration, the greater is the altitude of the last spiral before all the propellant is used up. This means that more energy is expended in raising propellant in the gravity field about the earth with low

accelerations than with high accelerations. The effect is to increase the propellant requirement and thus reduce payload. This is shown in Fig. 8 where the mass ratio, which is the pounds of empty weight at completion of the mission for each pound of initial weight, is plotted as a function of initial acceleration for the various interplanetary missions discussed previously. A low mass ratio corresponds to low payload. The reduction in mass ratio with decrease in acceleration is shown to be quite important especially for the more difficult missions. For the minimum-energy probe mission, the mass ratio is reduced 39% in reducing acceleration from 1 to 10^{-3} *g*. The corresponding reductions are 47 and 60%, respectively, for the minimum-energy round trip and the excess-energy probe missions. The mass ratio is reduced by 77% for the swing-around trip, and 81% for the excess-energy round trip to Mars. These results are for a specific impulse of 940 lb/(lb/sec) corresponding to a temperature of 4500 F and gas pressure of 10 psia. From this figure then, the accelerations chosen should be between 10^{-1} and 1 *g* for the best performance unless some means can be found to minimize the loss in energy due to lifting the propellant to high altitudes. Methods of programming thrust, which will not be treated herein, can be devised which consume most of the propellant at lower orbital altitudes. In this case, acceptable accelerations should be reduced by a factor of 10 or more.

Conceptual Space Nuclear Rocket

For all the missions except the excess energy round trip mission, accelerations of 10^{-2} *g* give mass ratios greater than 0.10, indicating reasonable vehicles which require no staging. If proper thrust programming is used to minimize the propellant gravity loss, then accelerations of 10^{-2} *g* can be used which require powers varying from 5 to 100 mw for initial spacecraft weights ranging from 25,000 to 500,000 lb.

Consideration of the power requirements for both space and booster nuclear rocket missions leads to the conclusion that perhaps the space nuclear rocket should be developed first since the low power requirements would greatly simplify the task. The emphasis in this case would be the attainment of high temperature and the solution of problems concerned with operation and control of high-temperature reactors. As experience is gained, higher power reactors could be attempted and perhaps used as second stages of larger chemical boosters to gain experience. Finally, the development of a high-power, high-temperature rocket for booster application should be undertaken. The use of nuclear energy for both the boost phase and the space phase of an interplanetary flight would compound the advantage over chemical systems.

The remainder of the paper is concerned with a low power space nuclear rocket concept and its estimated performance. It should be considered as the first of a series of reactors of increasing power leading to the ultimate high-power nuclear rocket booster. As the first reactor it would be used to determine the problems of attaining high performance and the problems of automatic control and operation of nuclear rocket powerplants.

In arriving at the configuration to be described, primary emphasis was placed on low weight, since the nuclear vehicle is to start from an earth orbit where it has been placed by a chemical booster. For a fixed chemical rocket booster payload capability, the lighter the powerplant of the boosted space vehicle, the greater is the allowable propellant and payload weight. No emphasis was placed on thrust level since space missions starting from orbit do not require large thrusts. Stress was placed, however, on high-temperature operation so that the specific impulse would be as high as possible. In addition, the low thrust requirement permits operation at low pressure, which further increases specific impulse due to dissociation. It was presumed that the NASA 1.5 million lb thrust engine would be available, and could place 25,000 lb in orbit at the time the nuclear rocket was ready for test.

System Description

Reactor — Assuming that chemical boosters will place the nuclear space rocket into an orbit about the earth, the reactor should be designed to be of minimum weight so that the payload and propellant weight can be maximized. In addition, the reactor should be constructed of materials which will produce the highest possible hydrogen temperatures.

Two principles are important in reduction of reactor weight. First, the total flow-passage volume should be as small as possible so that the reactor will include as much moderator as possible. A flow area equal to 10% of the reactor core frontal area was chosen for this reason. Such a low-flow area will make the hydrogen flow rate small. However, since large thrusts are not required for flights which start from orbit, this feature will not greatly penalize the nuclear rocket performance. The second principle is that the moderator should be chosen on the basis of ability to produce minimum reactor size and to operate at as high a temperature as possible. Of the available moderating materials, beryllium oxide and graphite are the most capable of operating at high temperature. Beryllium oxide can operate at temperatures approaching 4000 F, while graphite can operate at temperatures of 5000 F. Beryllium oxide is a much better moderating material than graphite and, therefore, will yield reactors of lighter weight. In addition, beryllium oxide is completely compatible with hydrogen up to temperatures of 4000 F, whereas, graphite is not compatible with hydrogen unless it is protected by a coating or an inhibitor contained in the hydrogen. It was, therefore, decided to use beryllium oxide as the moderating material for the conceptual design study.

Regarding the attainment of high hydrogen temperature, the materials which contain the fissioning uranium and provide heat-transfer surface to heat the hydrogen must have the highest possible operating temperature and be compatible with both hydrogen and the uranium-bearing compound which they contain. Of all the metals, tungsten has the highest melting point, has the highest strength at elevated temperatures, and is compatible with both hydrogen and uranium dioxide. This makes tungsten the best and most obvious choice among metals. The fuel elements, which are defined as those parts of the reactor which contain the fissioning uranium and supply heat-transfer surface for heating the hydrogen, are assumed to be made of tungsten containing uranium dioxide.

Such a fuel element is expected to be capable of operating at temperatures up to 5000 F. For comparison, the melting point of tungsten is 6100 F.

Tungsten is quite difficult to work because of its high-temperature strength and brittleness. Therefore, the fuel elements should be designed to be as simple as possible from the point of view of fabrication. Flat or slightly curved plates made by powder metallurgy techniques are perhaps the simplest shapes to make.

Another difficulty connected with tungsten is its high neutron capture cross-section, which means it competes with uranium for the neutrons generated in the reactor. Natural tungsten consists of a series of isotopes: 180, 182, 183, 184, and 186. Of these only tungsten-184 has a suitably low capture cross-section. Tungsten-184 can be separated from natural tungsten. It would be a costly process, but the savings in uranium over a natural tungsten reactor may pay for the separation of the tungsten isotope.

Using the above considerations, the reactor concept shown in Fig. 9 was evolved. The reactor is a right circular cylinder 28 in. in diameter and 36 in. long with fuel elements and moderator arranged in concentric annuli. The fuel elements are arranged in two of the concentric annuli, separated by beryllium oxide moderator regions. The nuclear computations were carried out using the techniques presented in Miser, et al.[2]

The fuel element design is shown in the inset. It consists of a series of slightly curved, uranium-dioxide-bearing, tungsten-184 plates 0.034 in. thick separated by a gap of 0.048 in. Each plate is about 2 in. long in the flow direction. The fuel elements have ears which fit into slots in the two side plates. The total height or thickness of the entire element is 0.459 in. A possible fabrication process consists of a cold pressing of the tungsten-184 and uranium dioxide powders in the ratio of 20% uranium dioxide to 80% tungsten by volume. This would be followed by a hot-sintering operation. The proper number of these elements are arranged to complete each annular ring in the reactor.

The moderator is considered to be made of hot-pressed blocks of beryllium oxide arranged between the fuel element rings. Cooling of the moderator is provided by the flow of hydrogen in the outer passage of the fuel element. The hydrogen temperature in the outer passages is kept low enough by proper choice of the outer passage height so that the beryllium oxide never exceeds 4000 F.

A 6-in. thick beryllium reflector, used to reduce the required core size, surrounds the outer fuel annulus. Beryllium metal should be used because its nuclear characteristics are slightly better than those of beryllium oxide and it has a lower density. Its relatively low temperature limit is not prohibitive for use as a reflector since the reflector is cooled by very low-temperature hydrogen. The reflector at the front end of the reactor is 6-in. thick beryllium. The rear reflector is 6-in. beryllium oxide since it is at the high-temperature end of the reactor where it is more difficult to cool. The overall weight of the reactor is 1900 lb.

Powerplant — The reactor core and its control rod structure and drive mechanisms are enclosed in a cylindrical pressure shell as shown in Fig. 10. The cooled exhaust nozzle is attached to one end of the pressure shell. Reactor control is provided by a series of neutron absorbing rods distributed uniformly through the moderator regions to provide a minimum of neutron flux distortion. They are operated by small electric motors located on a bulkhead in the pressure vessel. The gamma heating is removed from the control rod mechanisms by the hydrogen before it flows through the reactor. Electrical power is supplied by solar cells. Hydrogen, pressurized in a storage tank, is supplied to the powerplant at a pressure of about 35 psia. The liquid first flows through the regeneratively cooled stainless-steel nozzle walls. It then passes through passages in the reflector to remove the heat generated by neutrons and gamma radiation as they pass through the beryllium. At this stage the hydrogen is completely vaporized and is at a temperature of about 200 R. It then is used to cool the control rod mechanisms and passes into the reactor at a pressure of 25 psia. The hydrogen first cools the forward beryllium reflector, then is heated as it passes over the surfaces of the tungsten fuel elements. At the exit the hydrogen, heated to 4500 F, is at a total pressure of 10 psia due to accumulated pressure drops. It is ejected through the nozzle with a velocity coefficient of 0.96 to produce thrust. The nozzle has an area ratio of 50/1 which produces a specific impulse of 940 at the hydrogen pressure and temperature. The hydrogen flow rate

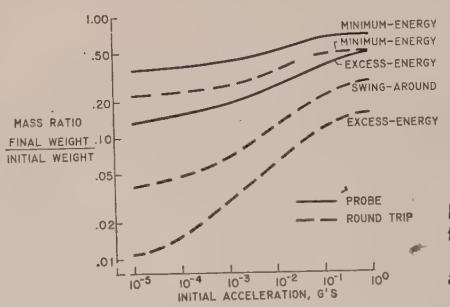
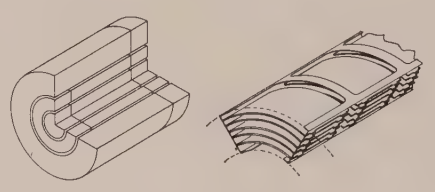


Fig. 8 — Mass ratios for nuclear rockets (300-mile initial orbit altitude, 4500 F temperature)



REACTOR DIAM, IN. 28
REACTOR LENGTH, IN. 36
REACTOR WEIGHT, LB. 1900

FUEL ELEMENT HEIGHT, IN. 0.458
GAP HEIGHT, IN. 0.048
PLATE THICKNESS, IN. 0.034
UO_2/W RATIO (VOL) 0.2-0.3

Fig. 9 — Space nuclear rocket reactor

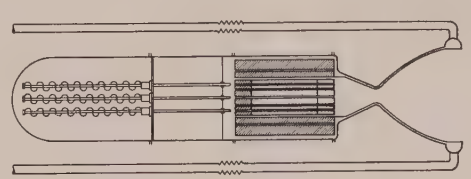
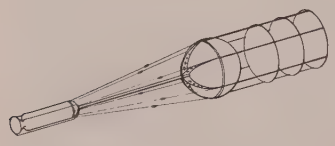


Fig. 10 — Powerplant cross-section

[2] "Computer Program for Solving Nine-Group Diffusion Equations for Cylindrical Reactors," by James W. Miser, Robert E. Hyland, and Daniel Fieno. NASA Memo 12-24-58E, 1959.

Table 1 — Powerplant Performance Characteristics

Reactor Inlet Hydrogen Pressure, psia	25
Reactor Exit Hydrogen Pressure, psia	10
Reactor Inlet Hydrogen Temperature, F	−260
Reactor Exit Hydrogen Temperature, F	4500
Maximum Fuel Element Temperature, F	5000
Hydrogen Flow Rate, lb/sec	0.35
Specific Impulse, lb/(lb/sec)	940
Thrust, lb	330
Reactor Power, megawatts (1 mw = 1300 hp)	6.7
Total Powerplant Weight, lb	2400



TOTAL POWERPLANT WEIGHT, LB	2,400
TANK WEIGHT, LB	700
STRUCTURE WEIGHT, LB	500
PROPELLANT WEIGHT, LB	12,100
PAYLOAD WEIGHT, LB	9,300
TOTAL WEIGHT IN INITIAL ORBIT, LB	25,000

Fig. 11 — Nuclear rocket space vehicle, minimum-energy Mars probe

is 0.35 lb/sec which produces a thrust of 330 lb. The maximum fuel element temperature is 5000 F, while the maximum moderator temperature is 4000 F. The reactor power is 6.7 mw. The additional weight of the control rod mechanism, nozzle, pressure shell, solar batteries, and guidance systems is assumed to be 500 lb. The total powerplant weight is then 2400 lb. The powerplant performance characteristics are summarized in Table 1.

Vehicle — Fig. 11 is an illustration of a nuclear rocket spacecraft using the powerplant described. The total weight of this vehicle on the earth's surface is 25,000 lb. It is boosted into a 300-mile orbit by a chemical rocket booster as described previously. The hydrogen in liquid form is stored in a spherical tank at a pressure of 35 psia. The tank is assumed for weight calculation purposes, to be constructed of 0.020-in. stainless steel and is pressure stabilized. The thrust forces are transmitted from the engine to the tank by means of the structure shown. The payload and electronic equipment are placed on the tank opposite the powerplant to give the greatest possible protection from the reactor radiations. No shielding is required to protect the liquid hydrogen if the angle subtended by the tank is less than 20 deg, since less than 1% would be evaporated as a result of reactor operation. The shielding required to prevent evaporation would weigh more than the propellant evaporated. This particular drawing shows the dimensions of the hydrogen tank for the minimum-energy Mars probe mission. For more difficult missions the payload would shrink and the tank volume would increase to accommodate the larger propellant requirement. For long interplanetary round-trip flights, liquid hydrogen must be stored in the presence of solar radiation. The hydrogen is protected during the coasting phases by means of the three solar radiation reflectors located in front of the tank as shown. They are very lightweight since they are fabricated of thin plastic sheet which has been gold plated. During the coast phase the spacecraft is oriented so that it always points toward the sun to make the solar radiation shields effective.

Performance

Calculations were made to indicate the performance that would be possible using the conceptual engine design (Figs. 9 and 10, Table 1) discussed in this paper. The performance is, therefore, not optimized, but serves to illustrate what can be expected of this particular powerplant for a variety of applications. Its performance is also compared with chemical and nuclear electric rockets.

Methods and Assumptions — The performance figures quoted in support of the foregoing design selection and development philosophy have been calculated by what are believed to be conventional methods. The simplifying assumptions which have been introduced to reduce the complexity of computation (that is, planets represented by central force fields and the solar system represented by circular, coplanar planetary orbits in a central force field), have been chosen to be consistent with the conceptual nature of the study. A brief outline of the analysis method and the principal assumptions follows.

Since the primary intention of the study was to investigate a space application, relatively little attention was paid to the surface-to-orbit trajectory. It was assumed that a multistage booster designed around the NASA 1,500,000-lb thrust chemical rocket engine would place 25,000 lb into a 300-mile earth orbit.

Missions which start from an earth orbit have been shown to be practical with thrust levels several orders of magnitude less than the initial weight in orbit. Analysis of low-thrust trajectories requires use of numerical techniques. The differential equations of motion which must be solved are based on the assumption of a point mass in a central force field with constant thrust parallel to the velocity vector. These equations have been solved in general form using a Runge-Kutta numerical integration method, and the results, although unpublished, were available to the authors at the time of the subject study.

For near-earth missions, such as orbit changes or lunar probes, the trajectory is assumed to consist of a constant-thrust propulsion period followed by a coast period to destination. When a new circular orbit is established at apogee (highest point of elliptic path), the required mass ratio has been estimated assuming impulsive velocity change. The lunar probe energy requirement was assumed equal to that for escape from the earth's gravitational field.

Space missions have been treated in essentially the same manner, but with the difference that the propulsion and coast periods are in different gravitational fields. The propulsion period during which the vehicle attains the desired hyperbolic velocity is assumed to be within the planet's field, while the coast period is assumed to be in the sun's field. The planet's orbits about the sun are assumed to be circular and coplanar. For the round trips illustrated, the spacecraft is considered to decelerate into Mars or earth orbit by means of reverse thrust. The orbits about Mars and earth upon return from Mars are assumed to be very eccentric elliptic orbits with minimum altitudes of 1000 miles at Mars and 300 miles at earth. The ratio of velocity at minimum altitude to circular velocity at that altitude is chosen to be 1.342.

This assumption, of course, defines the elliptical orbit.

The choice of particular trajectories to illustrate the earth-Mars missions was quite arbitrary. For the Mars probes, the excess-energy path chosen for comparison with the minimum-energy is characterized by an 80-day coast time and a tangent departure from earth's orbit about the sun. When the probe reaches Mars, earth and Mars will be at their closest approach, known as planetary opposition, thereby minimizing communication distance.

For the excess-energy round-trip missions, two illustrations are selected. One is called the "swing around Mars" mission where the vehicle leaves the earth's orbit about the sun at an angle which gives an elliptical coast time of 226 days to intercept Mars. At Mars the vehicle is deflected by the local gravitational field without any applied propulsive power. The vehicle then coasts back to earth's orbit on a 282-day path. Thrust is applied and the vehicle is trapped in an elliptical orbit about the earth with a perigee of 300 miles. The other excess-energy round trip is selected from another family of trajectories which are tangent to Mars' orbit about the sun. This is compared to the minimum-energy trip which is tangent to both earth and Mars orbits about the sun. The coast time is 270 days for the excess-energy trip compared to 259 days for the minimum-energy path. The wait time is 30 days compared to 455 days. Thus, by a proper choice of trajectories it is possible to reduce the total trip time from 973 to 570 days.

The latter excess-energy round trip was chosen as representative of low-wait-time paths, which, if symmetric, will also be characterized by planetary opposition during the wait time. However, the energy requirements for this round trip are too high to be compatible with an initial weight in orbit of 25,000 lb. In order to accomplish this mission more propellant is required and hence the initial weight must be increased.

Nuclear Rocket Performance — The performance of rocket vehicles using the nuclear rocket powerplant described in this paper for near-earth missions is shown in Fig. 12. The vehicles are boosted into a 300-mile orbit by a chemical booster system. The payload which can be carried on various missions is plotted as a function of gas temperature. The three missions consider transfers from a 300-mile circular orbit into (1) an elliptic orbit with a 50,000-mile apogee for exploring space environment and its effect on nuclear spacecraft in the immediate vicinity of the earth; (2) a 22,300-mile circular (stationary) orbit for communications or weather surveillance; (3) a trajectory which will carry it to the vicinity of the moon (lunar probe). In computing the performance at temperatures lower or greater than 4500 F, it was assumed that only the hydrogen flow rate and thrust would be affected, and that the powerplant weight would remain unchanged at 2400 lb. For gas temperatures as low as 1500 F payloads in the range of 6000–8000 lb can be carried on the three missions. This represents about one-fourth to one-third of the vehicle weight. Increasing the temperature to 3000 and 4000 F increases the payloads to the range of 10,000 and 12,000 lb, respectively. Operation at 6000 F would result in payloads in the range of 14,000 lb which is more than 50% of the initial vehicle weight. The examples shown are representative of

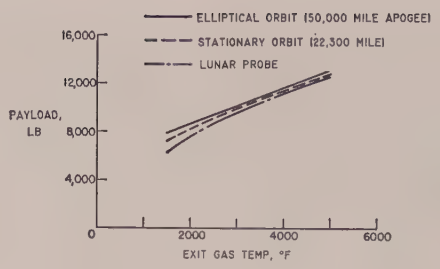


Fig. 12 — Nuclear rocket payloads for near-earth missions (25,000 lb initial weight in orbit, 300-mile orbit altitude)

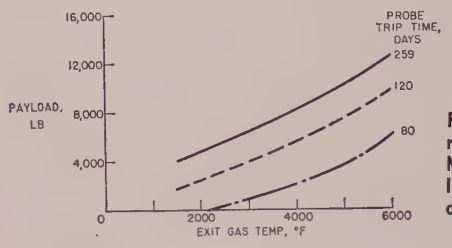


Fig. 13 — Nuclear rocket payloads for Mars missions (25,000 lb initial weight in orbit, 300-mile orbit altitude)

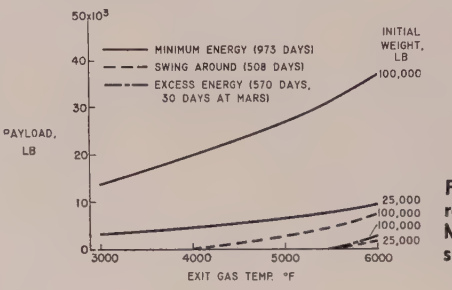


Fig. 14 — Nuclear rocket payload for Mars round-trip missions (300-mile initial orbit altitude)

the more difficult near-earth missions such a vehicle may be asked to fulfill. Should lesser payloads be sufficient for the purpose, it would not be necessary to use an initial weight in orbit as high as 25,000 lb.

The performance of rocket vehicles using the same nuclear rocket powerplant is shown in Fig. 13 for several Mars probe missions. Again, payload is plotted as a function of temperature. Three one-way trajectories are shown: 259-day coast, 120-day coast, and 80-day coast. The effect of temperature is more pronounced in this figure since more propellant is required for these more difficult missions. At 1500 F, for example, the 80-day trip is not possible at all. A temperature of at least 2000 F must be attained to have any payload capability at all for this mission. At 1500 F about 2000 and 4000 lb can be placed in the vicinity of Mars for the 120- and 259-day trips respectively. At 4500 F about 2700, 6500, and 9300 lb can be carried for the three illustrated trip times. Further large improvements could be made if the reactor could be built to operate at temperatures approaching 6000 F.

Fig. 14 shows the payload carrying ability as a function of temperature for round-trip Mars missions. The missions considered are the 973-day minimum-energy round trip, the 508-day swing-around trip, and a 570-day excess-energy round trip. Initial weights of 25,000 and 100,000 lb are assumed for each of the missions except the 570-day excess-energy trip which is not possible at the lower weight except for temperatures higher than 6000 F. All the

trips are powered by the 6.7-mw powerplant discussed in this paper.

At a gas temperature of 4500 F about 5500 lb can be carried to Mars and returned to an orbit about the earth along a minimum-energy path with an initial weight in a 300-mile orbit about earth of 25,000 lb. With 100,000 lb initial weight in orbit, the payload is increased to 23,000 lb. In the case of the 508-day swing-around trip it is not possible to achieve the desired mission with the 6.7 mw reactor with an initial weight of 25,000 lb unless the temperature is at least 5500 F. This temperature limit can be lowered if staging were employed, and if empty tanks and unnecessary structure were ejected before each powered maneuver. With an initial weight of 100,000 lb the payload is 1500 lb at a temperature of 4500 F. This payload can be increased to 7500 lb at a temperature of 6000 F.

The excess-energy 570-day round-trip mission, with an initial weight of 100,000 lb has no payload for the 6.7-mw powerplant unless the temperature is greater than 5500 F. At 6000 F the payload would be 3000 lb. The performance would be improved considerably by using larger engines and by the use of staging, however this is beyond the scope of this paper.

In all of the round-trip missions very substantial benefits occur by the use of temperatures approaching 6000 F. Temperatures of this magnitude may be attainable by the use of ceramic materials such as hafnium and tantalum carbides which have melting points of about 7000 F.

Comparison with Other Systems — Fig. 15 indicates the performance possible using the 6.7-mw nuclear rocket when compared to chemical rockets and nuclear turboelectric ion propulsion systems. The Mars probe payload is plotted as a function of trip time for the three systems. The nuclear rocket performance is shown for temperatures of 1500, 3000, 4500, and 6000 F. The powerplant used for each operating point illustrated in the figure is the one described in this paper (Figs. 9 and 10, Table 1). Actually, the performance could be improved over that shown if the optimum engine size were used for each operating point.

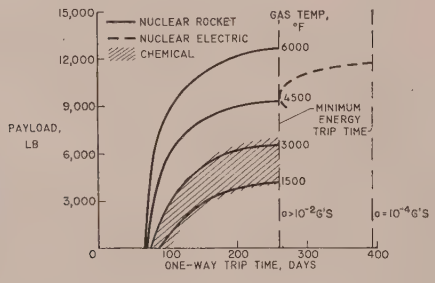


Fig. 15 — Comparison of systems for Mars probe mission (25,000 lb initial weight in orbit, 300-mile orbit altitude)

Chemical rocket performance is indicated by a band. The lower edge of the band corresponds to conventional gasoline-oxygen propellants, while the upper edge is for high-energy propellants such as hydrogen-oxygen. The lower conventional chemical rocket performance compares very well with the 1500 F nuclear rocket, while the high-energy chemical rocket compares favorably with the 3000 F nuclear rocket.

The nuclear electric system performance was obtained by using powerplant weights which could be achieved by the use of a nuclear sodium vapor turboelectric ion propulsion machine.[3] The turbine inlet temperature is assumed to be 2000 F. The powerplant weight per electrical kilowatt for this system is assumed to vary from 80 lb/kw at a power level of 10 kw to 27 lb/kw at 100 kw and 15 lb/kw for powers of 1000 kw and greater. The efficiency of converting electrical to kinetic energy in the ion jet is assumed to be 0.67. The ion propulsion engine is operated at a specific impulse of 5000 lb/(lb/sec) which is found to be the best value for an acceleration of 10^{-4} g and the minimum-energy Mars trip.

The nuclear electric system has longer trip times than the nuclear rocket or chemical systems. This is because of the low acceleration or high powerplant weight of ion propulsion system. The thrust is of the order of 10^{-4} times the initial weight so that a great deal of time is required to get up to speed. If more speed is desired to reduce the coasting portion of the flight, still more propulsion time is required. A point is reached where the savings in coast time is more than offset by the increase in propulsion time and the curve doubles back as shown. This results in a minimum time for the nuclear electric system of 260 days. At this trip time the payload of the nuclear rocket and nuclear electric systems are about the same. The nuclear rocket has the feature that payload can be sacrificed to shorten trip times. For example, in the case of the 4500 F nuclear rocket, a 50% reduction in payload permits a trip time reduction from 259 to 100 days. If payloads of 2000 lb are all that is necessary, the trip time could be reduced to about 76 days.

The performance of the three propulsion systems for round trip Mars missions is tabulated in Table 2. All the missions start from a 300-mile orbit about the earth. The powerplant for all the nuclear rocket vehicles except in the case of the 500,000-lb initial weight, is the 6.7-mw powerplant described in this report. At 500,000 lb the powerplant is scaled up to produce 35,000 lb of thrust needed for near optimum acceleration. The performance of the nuclear rocket could be improved by optimizing the engine design for each mission and by the use of staging.

Table 2 — Comparison of Systems for Round-Trip Mars Missions
(Initial orbit altitude, 300 miles)

	Propulsion System	I, $\frac{\text{lb}}{\text{lb/sec}}$	Temperature, F	Payload, lb	Total Trip Time, days
Minimum-Energy Mars	Chemical	400	—	1,800	973
Round Trip (Initial	Nuclear rocket	940	4500	5,500	973
weight in orbit,	Nuclear rocket	1330	6000	9,300	973
25,000 lb)	Nuclear electric	5000	—	10,200	1099
Swing-Around Mars	Chemical	400	—	0	—
(Initial weight in	Nuclear rocket	940	4500	1,100	508
orbit, 100,000 lb)	Nuclear rocket	1330	6000	7,400	508
	Nuclear electric	5000	—	27,800	803
Excess-Energy Mars	Chemical	400	—	0	—
Round Trip (30 days	Nuclear rocket	940	4500	7,500	570
at Mars; initial	Nuclear rocket	1300	6000	49,000	570
weight in orbit,	Nuclear electric	5000	—	0	—
500,000 lb)					

[3] "20,000-Kilowatt Nuclear Turboelectric Power Supply for Manned Space Vehicles," by Robert E. English, Henry O. Slone, Daniel T. Bernatowicz, Elmer H. Davison, and Seymour Lieblein. NASA Memo 2-20-59E, 1959.

The nuclear electric powerplant size is chosen to give the best performance when operating at a specific impulse of 5000 lb/(lb/sec) and at an acceleration of 10^{-4} g. Only chemical rockets with a specific impulse of 400 lb/(lb/sec) were considered for comparison on the round-trip missions.

For the minimum-energy Mars round trip the initial weight in orbit is 25,000 lb. The nuclear rocket, operating at a temperature of 4500 F, shows a payload about one-half of the nuclear electric system. At 6000 F the nuclear rocket has about 90% of the payload of the electric system. The nuclear electric system however, has a total trip time which is about 13% greater than the nuclear rocket. The chemical rocket has a payload which is only 1800 lb.

For the Mars swing-around mission, which requires more energy and consequently more propellant, the initial gross weight was increased to 100,000 lb. Because of the fact that only one engine of the type described in this paper was used for this large initial weight, the initial acceleration is less than 10^{-2} g for the nuclear rocket which causes large gravity losses (Fig. 8) and consequently the performance is not as good as it should be if the engine were to be designed for this mission. Even at a temperature of 6000 F the nuclear rocket shows a payload that is only about 27% of the nuclear electric system. The nuclear electric system, however, requires about 60% more time to accomplish this mission. The chemical system cannot accomplish this mission.

The final round-trip Mars mission shown is the 570-day excess-energy mission with a 30-day wait in orbit about Mars. The initial weight is increased to 500,000 lb. The nuclear rocket in this case uses a scaled up engine which weighs 17,000 lb and has a thrust of 35,000 lb. The importance of obtaining high temperature for the more difficult missions can readily be seen in this as well as the previous mission. The nuclear electric system suffers in the case of excess-energy missions which have relatively short waiting times. In order to achieve the waiting time of 30 days in Mars orbit the time for deceleration and acceleration must be added to the waiting time. As shown in Fig. 7 it is not possible to accomplish this mission at accelerations less than 2×10^{-4} g. The nuclear electric system has accelerations less than this, so it cannot do this mission unless the waiting time is increased by at least 780 days. Trajectories for nuclear electric systems have not as yet been thoroughly investigated and there may be trajectories which enable it to accomplish a 570-day trip with a waiting time in orbit at Mars of 30 days. This mission, if it can be accomplished at all with the electric system, would require large amounts of excess energy with consequent reductions in payloads.

Concluding Remarks

It has been shown that large reactor powers are required for the booster application of nuclear rockets. The power requirements are reduced by about three orders of magnitude for even the most ambitious interplanetary missions, if the nuclear rocket spacecraft can be boosted into a low-altitude earth orbit by chemical boosters now under development. It may be wise, for this reason, to first develop the use of nuclear energy for rocket propulsion as a space engine which starts from a low-altitude earth orbit. As the technology of high-temperature, remotely operated hydrogen-cooled nuclear reactors advances, higher power powerplants should be built in steps of factors of 10 or so, until the power level required for boosting 500,000 lb into an earth orbit is attained. Such a series of development reactors could produce useful propulsion devices at each step of the way. First, space engines boosted in orbit by chemical boosters would be built for interplanetary missions. Next, powers adequate for second-stage systems would appear. And finally, the ultimate very large booster nuclear rocket systems could be developed to complete the application of heat-transfer-type nuclear rocket engines.

A particular space nuclear rocket engine concept, which is suggested as the first application of nuclear energy, is described in a fair amount of detail in this paper. This particular engine even though not optimized for each mission shows excellent performance, particularly at temperatures above 3000 F, for near-earth missions, lunar probes, Mars probes, and Mars round-trip missions when compared to the best chemical systems. At temperatures of 4500 F, a clear superiority is shown, which is increased by going to the ultimate temperature of 6000 R. The nuclear rocket system is competitive with the nuclear electric ion propulsion system on the basis of payload carrying ability for Mars probe missions. For all missions the nuclear rocket has a trip time advantage because of its higher acceleration compared to the nuclear electric systems. The nuclear electric system seems to be superior in payload carrying ability for the Mars swing-around mission which does not require propulsion at Mars. The trip time required for the electric system, however, is about 60% greater than for the nuclear rocket system due to long propulsions times required at the earth. The excess-energy Mars round-trip mission with a 30-day wait at Mars cannot be accomplished with a nuclear electric system with a 10^{-4} g acceleration. The nuclear rocket system can perform this mission with a total trip time of 570 days.

Very ambitious space missions (500,000-lb initial weight) require power levels of about 100 mw if they can be initiated from earth orbits. The nuclear rocket concept discussed in this paper operates at a power level of 6.7 mw and it would be a relatively simple task to redesign the reactor for the 100 mw power level.

Of course, the engine concept discussed in this paper is but one of many that could be conceived. However, all high specific impulse nuclear engines will have at least one thing in common, that is, they will require the development of materials like tungsten or the higher melting carbides to produce the high-temperature hydrogen. Further thought is being given to reactors using molten-uranium-bearing alloys or gaseous uranium compounds which can operate at temperatures beyond the highest material melting point. In these systems heat is transferred to gaseous hydrogen by direct contact with the liquid or gaseous uranium compounds.

Many other novel ways of utilizing more fully the potential of nuclear energy are under investigation, and perhaps one of these will provide a much superior system. Until then, we must be satisfied with the more or less conventional approach which was illustrated in this paper. Even with this approach, however, very attractive performance can be expected.

TRANSISTOR

THE TRANSISTOR SWITCHED ignition system may replace induction coil for applications where a premium can be paid for reliability. This paper describes the design features and advantages of this system, along with applications.

The new system can be used interchangeably with the present conventional induction coil. Using a slow coil (high-tension transformer) the system features:

1. Voltage output which is relatively flat throughout the engine speed range.
2. Reduced contact current giving unlimited electrical life and reliability of contacts.
3. Performance under fouling conditions at least equal to the present induction coil.
4. Voltage ceiling flexibility — the voltage can be raised to provide for possible future engine designs.*

A GENERAL TREND in automobile engine design developed during the past decade toward increasing maximum horsepower with minimum fuel consumption (Fig. 1). This trend to higher efficiency engines has emphasized the present ignition system limits and inadequacy as to its ability to fire fouled spark plugs and develop sufficient voltage for the higher compression ratio, higher speed engines.

Spark-plug fouling is a more critical factor in ignition system performance than ever before. The increasing compression ratios demand increased antiknock qualities in gasoline, as shown in Fig. 2. These increased antiknock qualities are largely built into fuels with increased amounts of lead bearing additives per unit volume. This increase in lead content has resulted in aggravating the spark-plug fouling problem. Field reports, driver complaints, and engineering studies emphasize the magnitude of the fouling problem in late model passenger cars operated by the motoring public.[1] (See Fig. 3.)

Higher efficiency engines have placed heavy demands on the voltage available from ignition systems. Greater spark gap lengths and higher compression ratios have raised voltage requirements throughout the speed range (Fig. 4). Higher speed engines and the wider use of 8-cyl engines emphasizes the need for adequate high-speed performance from the ignition system. It is not uncommon to have modern passenger-car engines rated at 4500 rpm and capable of attaining 6000 rpm or better, as opposed to approximately 3600 rpm rated and approximately 5000 rpm top of a few years ago.

Not New Problems

Inability to fire fouled spark plugs and poor high-speed performance of the conventional ignition system are not new problems.

The high-speed performance of the battery ignition system can be and has been improved through securing as high a primary inductance as possible with minimum decrease in current with increasing engine speeds. (See Appendix I.)

The following design devices have given partial relief:

1. Ballast Resistors
2. Increased Dwell Time
   a. Cam design
   b. Double breaker systems
3. Twelve-Volt Systems

All of these expedients have contributed in some degree to improve the high-speed performance of the conventional, high-voltage ignition system. However, a general increase of the overall voltage or energy at all speeds is dictated if the needs of very high compression ratios (12/1 and higher) and nonfouling ignition systems are to be met.

If we are to meet the demands of higher compression ratios and increased amount of lead, we must eliminate the inhibitions imposed upon design by the contacts (Appendix I).

Nonmechanical switches of the electronic tube variety have presented incompatible design considerations. In general, this type of switch presents the designer with one or more of the following problems:

1. Fragility.
2. Filament warmup time.
3. High plate voltage required (usually 300–500 volts).
4. Close system voltage regulation.
5. Poor or no output at low battery voltage (cold starting).
6. Complex circuitry.
7. Bulky system packaging.
8. High cost.

Power Transistor

The introduction of the power transistor appeared to offer an electronic switch overcoming many of the objections of tube types. The prospect of employing power transistors to switch the increased

* Paper presented at SAE Annual Meeting, Detroit, Jan. 14, 1960.

[1] "Spark Plug Fouling," by C. A. Hall, R. C. Beaubier, E. C. Marckwardt, and R. L. Courtney. Paper presented at SAE Annual Meeting, Detroit, January, 1957.

SWITCHED IGNITION SYSTEMS

G. E. Spaulding, Jr.
The Electric Autolite Co.

currents imposed by increased ignition demands was attractive, since such a system would:

1. Be rugged (no glass envelopes).
2. Eliminate warmup time (no filaments).
3. Operate directly at 6, 12, or 24 volts.
4. Operate over a wide range of battery voltages from cold starting to normal operating.
5. Permit miniaturized circuitry.

Even with these definite advantages, it was obvious that the circuitry would be more complex than conventional systems, and the cost would be inherently greater. However, the pressing need for systems capable of developing either ultra-high voltage or ultra-high energies at the spark plug dictated the exploitation of the solid electronic switch, the transistor.

Two types of transistors were available — silicon and germanium. Each has its own electrical, physical, and economic characteristics.

The silicon power transistor has temperature and voltage characteristics compatible to the environment of the conventional ignition circuit; on the other hand, their maximum rated current is about 5 amp and they are much more expensive than germanium. The current limitations indicated that more than one would be required in the high-energy, high-current systems contemplated.

The use of multiple silicon transistors would certainly preclude any commercial consideration in the immediate future due to their cost.

The germanium transistor, however, was enough

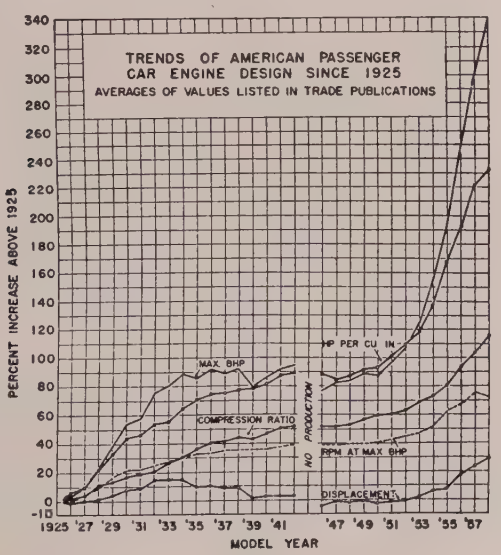


Fig. 1 — Trend toward higher horsepower, higher efficiency engines has accelerated during past decade (courtesy: Ethyl Corp.)

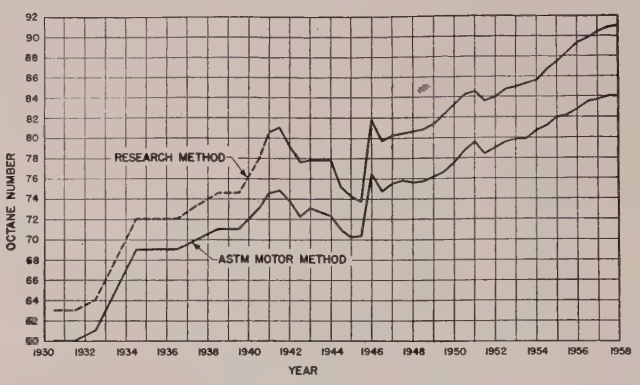


Fig. 2 — Higher efficiency engines demand higher octane fuel (courtesy: Ethyl Corp.)

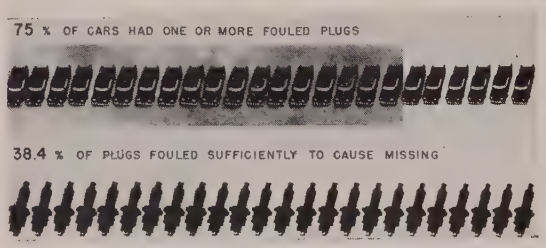


Fig. 3 — Seriousness of spark-plug fouling problem as shown in Autolite spark-plug surveys of four makers of 1954 and 1955 models

cheaper to offer commercial possibilities. This type of transistor is rated as high as 30 amp, but will not tolerate voltages generated in the primary of the ignition circuit, nor will it tolerate the ambient temperatures reached in some underhood applications.

One other electrical restriction imposed by both types of transistors should be noted. Only PNP types were available for ignition applications. This meant that development would be restricted to negative ground systems until NPN types of sufficient capacity were available.[2]

Since the majority of U. S. electrical systems are negative ground, this did not appear to offer any great commercial objections.

Research and Development Objectives

The decision was made in early 1956 to proceed with development work within the limitations imposed by available germanium power transistors.

It was felt that if silicon should become economically more attractive at any time, the basic circuitry and concepts could be readily adapted permitting a quick substitution.

This fundamental decision permitted research to proceed with an optimum of program flexibility.

The development problems then resolved into:

1. Design to meet temperature limitations.
2. Design for negative ground.
3. Design to prevent damage if inadvertently connected positive ground.
4. Design to develop sufficient voltage for starting at markedly reduced primary voltages (as low as 6 volts in a 12-volt system).
5. Design to meet severe vibration and rough handling encountered on engine applications.
6. Design as compact as possible.
7. Design for economy.
8. Design for simplicity so that repairs can be effected without sophisticated equipment or extensive electronic know-how.
9. Design to prevent destruction due to voltage induced in the primary.
10. Design so that the basic power supply could fire fouled spark plugs as well as, if not better than, conventional systems.
11. Design to permit adaptation to high-energy, nonfouling ignition systems.

All the above design criteria could be met in a relatively straightforward manner, except Items 9 and 10.

Induction Coil Versus "Slow Coil"

Reduction of back voltage in conventional ignition coils is impossible because they operate on the induction coil principle. A high-tension transformer, the "slow coil" of yesteryear, had to be considered if germanium transistors were to be used. (See Appendix II.)

The high-tension transformer was abandoned in the early 1920's in favor of the induction coil for reasons both economic and operational.

The induction coil was smaller, lighter, and cheaper to construct. In addition, the induction coil with its 1500–2500 cycle secondary wave front exhibited appreciably better ability to fire carbon-fouled spark plugs than the approximately 500-cycle transformer. This difference in performance under resistive load is shown in Fig. 5. (See Appendix III.)

On the other hand, the induction coil with its inherently high primary inductance and high primary induced voltage brought with it much more serious contact problems. These characteristics of the induction coil are capable of initiating and sustaining an appreciable arc at the separating contacts resulting in severe "burning" and metal transfer. (See Fig. 6.)

The early induction coils took this more or less in stride, since a relatively low primary current could provide adequate voltage for the low-compression engines of that day.

We have all observed, however, how the increased current requirements of modern ignition has seriously foreshortened contact life.

A return to the "slow coil" then would provide a measure of relief in contact life, permit the use of transistors, but threaten to reduce system efficiency firing fouled plugs.

Research Program

It was decided then to proceed with two concurrent programs:

1. The investigation and determination of the quantitative difference between the two coils under fouling conditions.
2. The application of transistors and the "slow coil" to those high-energy, nonfouling systems where the voltage buildup is isolated from the spark plug.

Comparative Fouling — "Slow Coil" Versus Induction Coil

Disregarding the testimony of the venerable "megohm test" (Appendix III) we proceeded to evaluate the relative fouling characterists of the two systems. (See Fig. 5).

Comparative tests were run in the dynamometer and on the road using accelerated fouling cycles.[3] These tests indicated that spark-plug miss occurred with both systems in the same general range of time and resistance values. Misses were recorded with both systems as early as 12 hr (336 miles) while in the same engine during the same tests, either system was repeatedly able to sustain firing without missing in excess of 84 hr (2348 miles). Shunt resistances of missing spark plugs were recorded at values varying from zero to 1.0 megohm with both systems; paradoxically, frequent readings as low as 0.2 megohms were made on plugs with no sign of missing.

The random results of these tests seem to indicate that other factors dominate the fouling picture.

Such variables as combustion-chamber temperature, fuel distribution, and variation of spark-plug rating within a given heat range can conspire to give wide differences between fouling characteristics between cylinders in any given engine.

Obviously, too, there is a dramatic difference between the effect of carbon fouling, from which the megohm test evolved, and lead fouling prevalent today.

---

[2] Of course, the PNP systems could be used with positive ground ignition systems if proper precautions are taken, such as two-wire distributor base plates with the base plate insulated from ground.

[3] On the dynamometer test the engine is cycled repeatedly at 500 rpm no load for 40 sec, to 1600 rpm, road load for 140 sec. Every 12 hr the engine is accelerated at full load to 3000 rpm and held there for 30 sec while each plug is checked for shunt resistance and missing as recorded on a Brush dual-channel recording oscillograph.

The road test is similar except the fouling cycle is from idle to 35 mph and the rating cycle is from 35 to 80 mph.

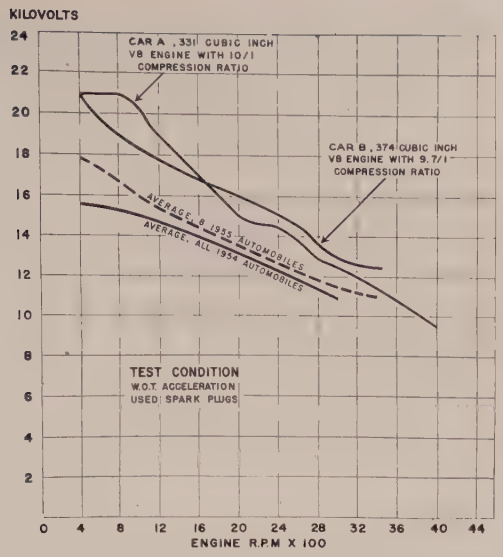

KILOVOLTS
CAR A, 331 CUBIC INCH V8 ENGINE WITH 10/1 COMPRESSION RATIO
CAR B, 374 CUBIC INCH V8 ENGINE WITH 9.7/1 COMPRESSION RATIO
AVERAGE, 8 1955 AUTOMOBILES
AVERAGE, ALL 1954 AUTOMOBILES
TEST CONDITION
W.O.T. ACCELERATION
USED SPARK PLUGS
ENGINE R.P.M. X 100

Fig. 4 — Increasing horsepower and efficiency of modern engines is reflected in increasing demands upon ignition system

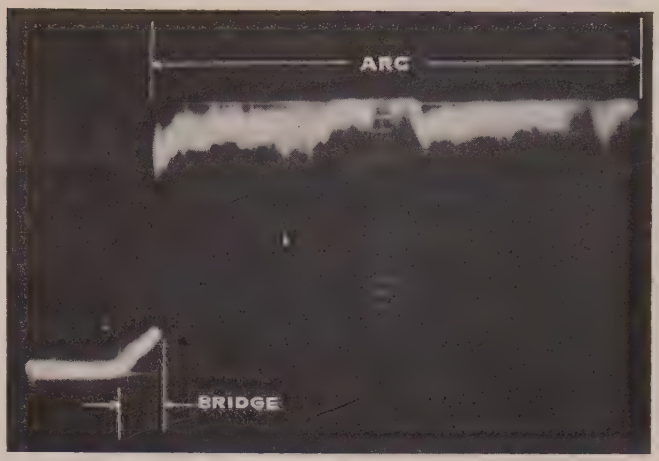


Fig. 6 — Bridge and arc between breaking contacts in induction coil system

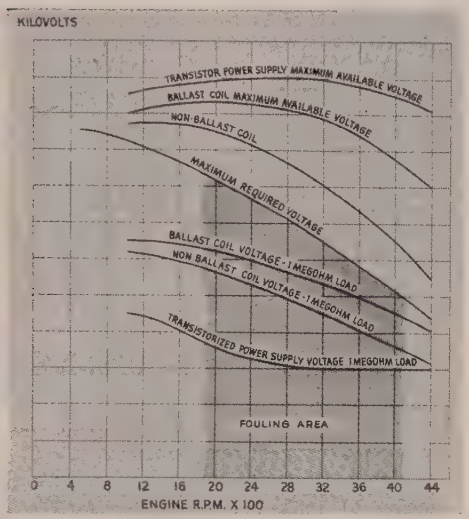

KILOVOLTS
TRANSISTOR POWER SUPPLY MAXIMUM AVAILABLE VOLTAGE
BALLAST COIL MAXIMUM AVAILABLE VOLTAGE
NON-BALLAST COIL
MAXIMUM REQUIRED VOLTAGE
BALLAST COIL VOLTAGE - 1 MEGOHM LOAD
NON BALLAST COIL VOLTAGE - 1 MEGOHM LOAD
TRANSISTORIZED POWER SUPPLY VOLTAGE 1 MEGOHM LOAD
FOULING AREA
ENGINE R.P.M. X 100

Fig. 5 — Comparative performance of transistor and conventional ignition systems

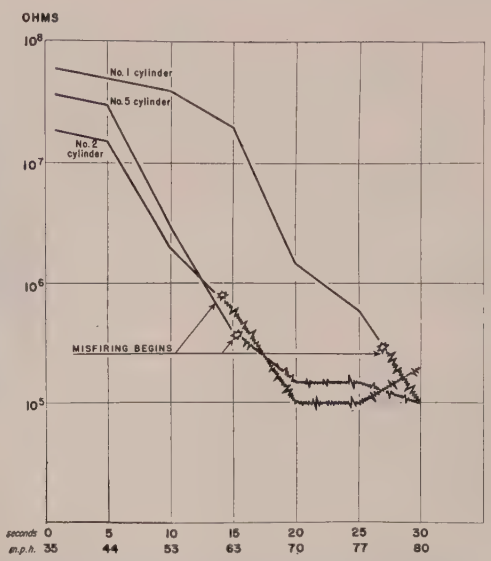


Fig. 7 — Spark-plug shunt resistance variation with vehicle speed under wide-open throttle acceleration

Under carbon fouling conditions, the spark plug will gradually lose resistance with a finite period of time, measurable in minutes or miles, required for it, say, to drop from 1.5 to 0.5 megohms — from a nonmissing to a missing condition. Under these conditions, with low-compression engines, it is desirable to maintain the "1 megohm" load performance of the ignition system as high as possible.[4]

Lead, on the other hand, is deposited on the spark-plug tip under low-temperature or part-throttle conditions. These deposits have a high electrical resistance — generally above 10 megohms — and will not significantly affect the ignition system at the operating condition under which they are formed. Most of these compounds, however, will react critically to temperature and will undergo chemical change or changes of state when they are subjected to higher temperatures.

Thus, a vehicle can operate for long periods of time under part-throttle conditions with no evidence of missing due to fouling. On the other hand, with the deposits present, when the vehicle's operating conditions are changed from low speed to high speed, the resistance of the deposits drop very fast. When this resistance which is shunting the gap reaches a certain critical value, the system will not fire.

Fig. 7 shows the rapid drop of resistance into the fouling range with increasing speed and load. Obviously, it is the rapid and steep drop of resistance at the critical temperature which results in the two systems exhibiting comparable fouling performance under lead fouling conditions. It is also obvious that the time until debilitating fouling of either ignition system is a function of the rate of lead deposition. Lead content of the fuel, spark-plug heat range, fuel distribution, and combustion-chamber temperature control this deposition rate and, therefore, are the dominant factors.

These controlled test results have been verified

[4] It should be noted, however, that even when carbon fouling is induced in modern high-compression, multicylinder engines, the required voltage at idle speeds is so high as to make the differences in time between the induction coil and the "slow coil" to the point of debilitating fouling insignificant. Both will foul in a matter of hours.

with many thousands of miles of vehicle testing under normal road and driving conditions.

Design Specifications

Assured that the "slow coil" could be expected to give at least equal performance under fouling, work proceeded to adapt the germanium transistor to a high-voltage ignition system.

Design specifications for the new system dictated:

1. Infinitely long electrical life of contacts.
2. Flat voltage output throughout the speed range.
3. Potential to lift the entire energy level or output voltage.

The reason for the first specification obviously was in line with demands for increased system reliability and longevity. The second specification would permit incorporation of series gaps to give a true antifouling system.

The third specification anticipated the needs of certain nonfouling ignition systems under development, as well as the possibility that 30 or 40 kv systems might be demanded in the future.[5]

Transistorized Ignition Circuit

Fig. 8 shows the circuit developed. The principle of operation is based upon the transistors ability to amplify current. When the contacts close, a small current flowing through the emitter-base circuit will switch on the transistor permitting a larger current flow through the emmiter-collector circuit. (See Fig. 9.) The current gain, or amplification factor, is a function of this transistor itself.

Fig. 10 shows the current carried by the electrical contacts compared to the current switched by the transistor. The amplification is about 30 to 1; con-

[5] Such ultra-high-voltage systems would, of course, demand redesign of all ignition components with the greater spacing and super dielectrics such tremendous potential would demand.

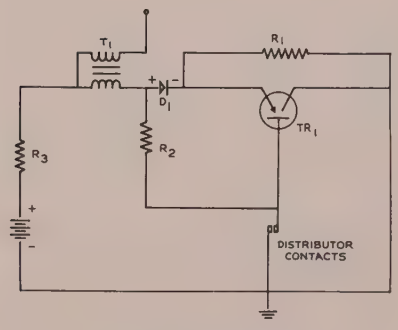


Fig. 8 — Transistor switching high-voltage ignition system

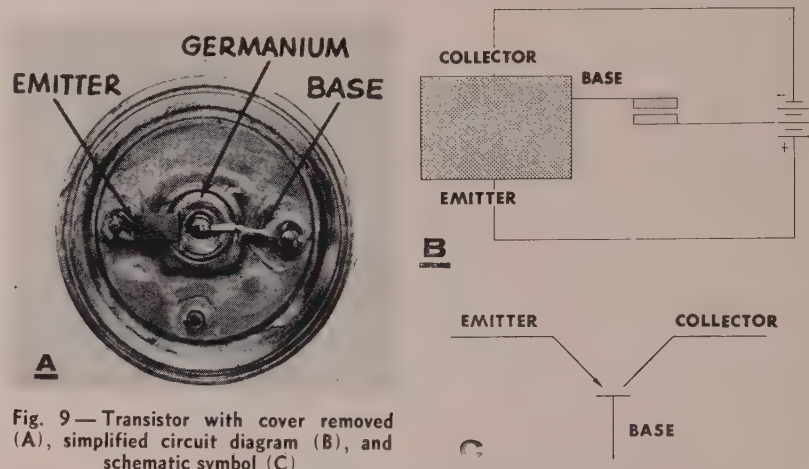


Fig. 9 — Transistor with cover removed (A), simplified circuit diagram (B), and schematic symbol (C)

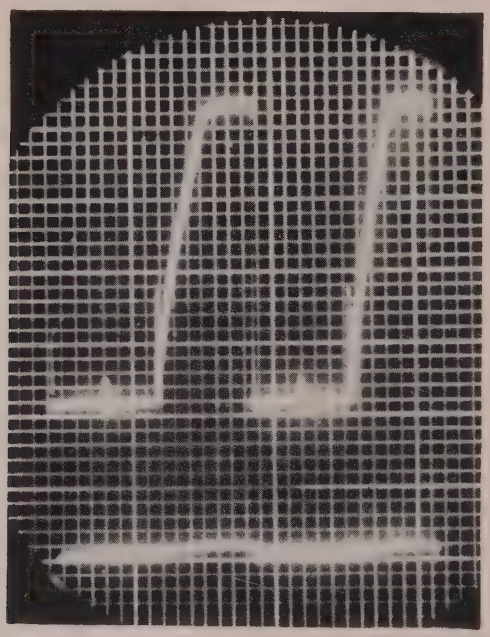


Fig. 10 — Current amplifying characteristic of transistor in ignition system. Collector current switched by transistor (top), base current switched by contacts (bottom). Horizontal axis: 13.7 millisec/in.; vertical axis: 5 amp/in.

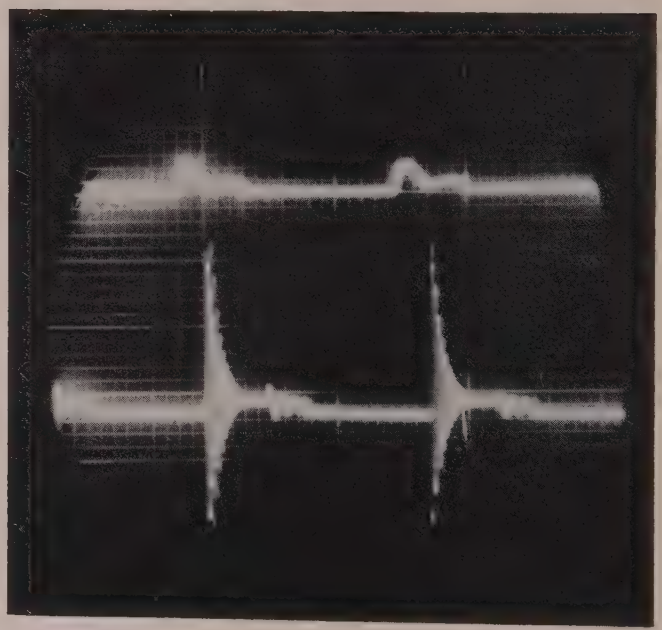


Fig. 11A — Primary wave forms of transistorized (top) and conventional (bottom) ignition systems compared. Dual beam oscilloscope. Horizontal axis: 13.7 millisec/in.; vertical axis: 500 volts/in.

tact current is about 0.250 amp, while transistor or system current is approximately 7.5 amp. The transistor in the circuit is a PNP transistor.

The relatively flat output voltage shown on Fig. 5 is obtained because of the fast rise of current to design value due to the relatively low primary inductance.

Diode D_1 provides a means of reverse biasing the emitter to base junction when the distributor contacts S_1 are open. This insures transistor cutoff even at high temperatures. Resistor R_1 allows a small current to flow thru D_1 continuously, thus there is a 0.5–0.75 volt drop across D_1. Since the transistor base is connected to the positive side of the diode thru R_2 and the emitter is directly connected to the negative side of the diode, when the contacts are open, the base is at a potential of 0.5–0.75 volts positive with respect to the emitter, which guarantees transistor cutoff. This action is enhanced by keeping R_2 as low as possible, but R_2 cannot be too low or it would allow too much current to flow through the contacts on make.

The diode has a peak inverse voltage rating of sufficient magnitude to protect the circuit in case it is connected with reverse polarity. During the dwell period when the contacts are closed, the transistor base is directly connected to the collector, allowing maximum current to flow. Transistor impedance when full on is on the order of 0.10 ohm. Initial current flow is limited by the inductance of the primary of the high-voltage transformer T_1 but finally by the ballast R_3.

Note that the traditional condenser across the primary contacts (S_1) has been eliminated.

Figs. 11A and 11B show the characteristic voltage wave forms induced in the primary and secondary of the two systems. Notice that the primary induced voltage of the transistorized system is on the order of 60 volts; about one-fifth the voltage induced in the induction coil primary. Also note that the frequency of the secondary voltage wave of the induction coil is about 1500 cycles as opposed to a 400-cycle wave in the "slow coil."

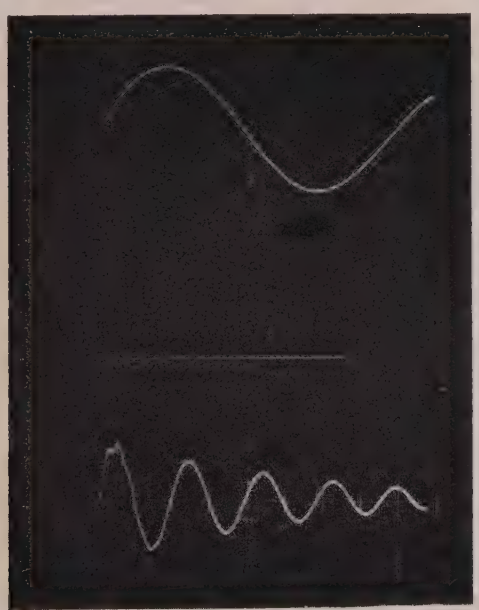


Fig. 11B — Secondary wave forms of transistorized (top) and conventional (bottom) ignition systems. Dual beam oscilloscope. Horizontal axis: 250 microsec/cm; vertical axis: 10 kv/cm

Design Features and Advantages

In summary, then, the first objective of the overall program has been realized. A high-voltage transistorized ignition system has been designed which can be used interchangeably with the present conventional induction coil (Fig. 12). The design fea-

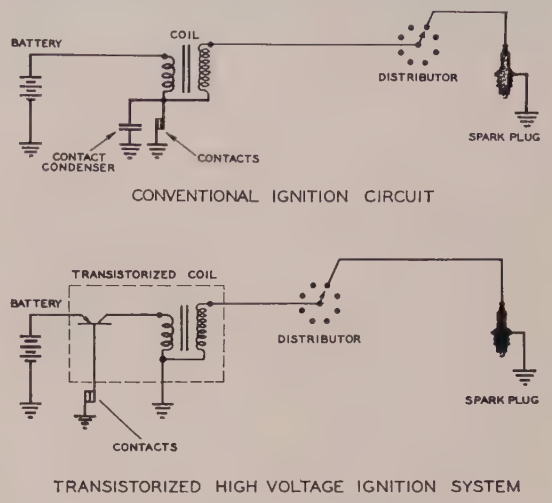


Fig. 12 — Schematic diagrams of conventional and transistor switched high-voltage ignition systems

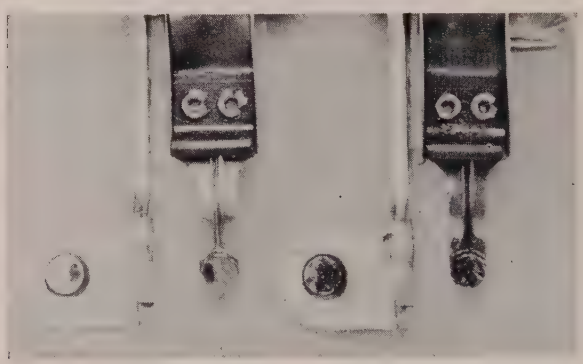


Fig. 13 — Direct comparison of contacts from systems after 44,000 miles of endurance tests (left — transistorized; right — standard)

Fig. 14 — Direct comparison of contacts from systems after accelerated cold tests at – 20 F (left — transistorized; right — standard)

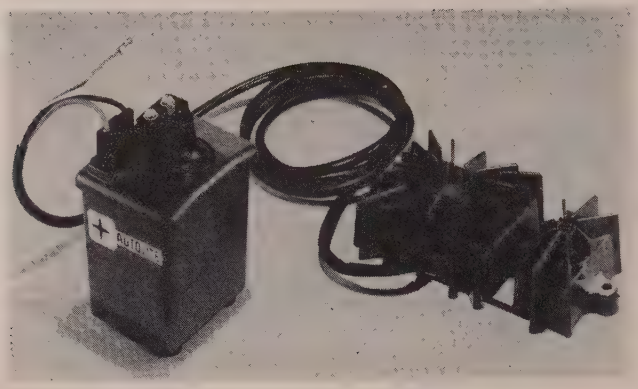


Fig. 15 — Prototype transistorized ignition system

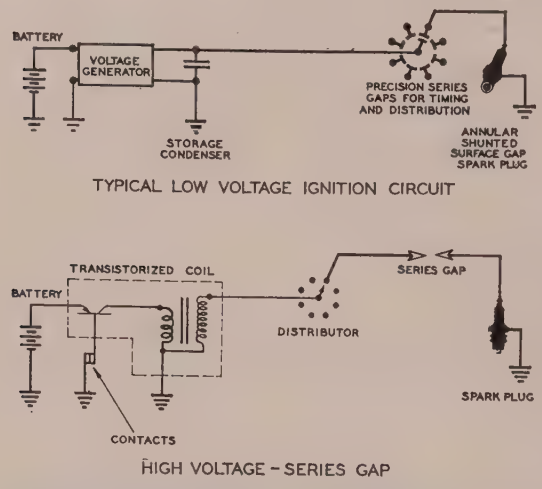


Fig. 16 — Schematic diagrams of high-voltage series gap (bottom) and low-voltage ignition (top) utilizing transistor switching technique

tures and advantages are as listed below:

1. Voltage output is relatively flat throughout the speed range (Fig. 5).
2. Contact current has been reduced giving unlimited electrical life and reliability of contacts. (See Figs. 10, 13–14.)
3. Performance under fouling conditions is at least equal to the present induction coil.
4. By appropriate changes in system parameters, the voltage ceiling can be lifted on demand to provide for possible future engine designs.
5. The coil package is rugged and may be located in ambients up to 350 F. (See Fig. 15.)
6. The separate transistor heat sink permits location in areas where the ambient does not exceed 180 F. (See Fig. 15.)
7. The system is designed for negative ground and is protected against inadvertent reversal of primary polarity.
8. The system is adequate for starting with battery voltages as low as 6 volts in a 12-volt system.

Immediate Applications

Obviously, the price of the package will be high when compared to the coil and condenser it will replace. There are, however, certain commercial applications which can justify the added cost when overall system reliability and longevity are prime economic factors. Some applications which appear to offer a ready market are:

1. Stationary engines providing pumping or stand-by power for the utilities industries.
2. Stationary engines providing power for heat pumps, airconditioning units, irrigation pumps, and the like.
3. Certain commercial fleet applications.
4. Certain off-the-road and agricultural equipment.

The gratifying results of the research and development program, coupled with indications from customer contact and market research that immediate markets existed, dictated that the high-voltage transistorized ignition system be released to production. Accordingly, a public announcement was made of these plans in February, 1959, and pilot line production of prototypes was begun.

The rigorous accelerated laboratory, environmental, and road tests disclosed several packaging and electrical inconsistencies that lowered the quality of the product below the high levels that had been established.

The physical configuration of the transformer itself resulted in unusually short flashover distances between the secondary and ground (the core). Initial plans to "pot" the coil in an epoxy had to be abandoned due to the extremely high voltages developed and the relatively frequent incidence of dielectric failures at these short flashover points.

Packaging in oil solved the dielectric problem but required redesign to hold the relatively massive transformer in position impervious to vibration encountered under field conditions. This change in concept and redesign dictated an entire new testing program to prove the ruggedness of the new unit. This proof testing is presently underway.

Several puzzling and frustrating system failures were traced to inadequate grounding of the collector. Originally No. 18 wire was used, but tests disclosed that the ohmic resistance of the length of lead used resulted in borderline conditions. When replaced with No. 12 wire transistor failures due to this cause have disappeared.

Variations in transistor amplification and breakdown voltage characteristics outside of specifications are not acceptable, necessitating establishment of elaborate inspection and quality control procedures.

All of these problems were, of course, anticipated; but the magnitude of some were underated in initial announcements of intended market introduction dates.

Future Applications

Several methods of combating the fouling problem through system design have intrigued designers for many years.

Although there is considerable difference in configuration of the various systems and their components, they all have at least two characteristics in common:

1. They involve charging a capacity and then discharging it across the spark gap. This results in an extremely high-frequency discharge, usually in excess of 1 megacycle. These systems will, then, cause a spark to occur at the plug even with very low values of shunt resistance.
2. They all require more energy input than is

available with conventional switching techniques.

As stated earlier, a second, and concurrent, phase of the master program involved the application of transistor switching concepts to the more promising "antifouling" systems.

Two basic types of systems are involved:

1. High voltage — auxiliary gap (see Fig. 16)
2. Low voltage — condenser discharge (see Fig. 16)

Although there have been reputable brand spark plugs on the market for years incorporating series gaps, their use has been restricted to special applications. The induction coil system is not capable of developing the voltage required by the combination auxiliary — spark-plug gap at starting and throughout the speed range (Fig. 17).

The transistorized system makes adequate voltage available as shown on Fig. 17. With the energy problem solved, research now revolves around development of an optimum series gap configuration.

The second system under development requires the use of special shunted surface gap spark plugs. A schematic of the system is shown on Fig. 16.

The U. S. Army has taken an active interest in this development work and is purchasing and evaluating several complete systems on Ordnance Contract DA-33-019-ORD-2740. In addition, Ordnance Contract DA-33-019-ORD-2748 has been issued covering a feasibility study on another extremely promising application of the system.

Summary

Transistor switched ignition systems are feasible. At the present time, programs are active covering:

1. Production prototype evaluation and development of a high-voltage transistor switched system intended to replace the conventional induction coil on applications where a premium can be paid for reliable operation.
2. Development of a high voltage — auxiliary gap system is now a possibility. This system would provide reliable, trouble-free performance on commercial applications where fouling (either lead or carbon) is a problem.
3. Development of low voltage — condenser discharge — surface gap system is underway. Prototype samples are being evaluated for several military and commercial applications.
4. Research is underway which is hoped will lead to an economical and practical way of switching the transistor without moving contacts of any kind. The low signal current switched by the contacts in the ignition system places such a configuration within the realm of possiblity.

Acknowledgments

I wish to acknowledge the contributions of the following, who are my associates at Electric Auto-Lite Co.: L. P. Atwell, J. F. Gage, K. E. Brandeburg, J. G. Naborowski, C. J. Dunne, D. B. Seem, G. Seidman, R. Burke, and W. Brumback.

APPENDIX I

In the conventional ignition system cycle, current flows through the primary winding of the coil creating a magnetic field when the distributor contacts are closed. When the contacts separate, this magnetic field collapses very quickly, inducing a high voltage in the secondary winding which charges the capacity associated with the secondary circuit to a high voltage. To achieve maximum output voltage from a given ignition coil, a maximum amount of energy must be stored in the magnetic field of the coil during the period the distributor contacts are closed. The amount of energy stored in this period depends on the primary current at break and the inductance of the primary winding. This relationship is expressed by the formula:

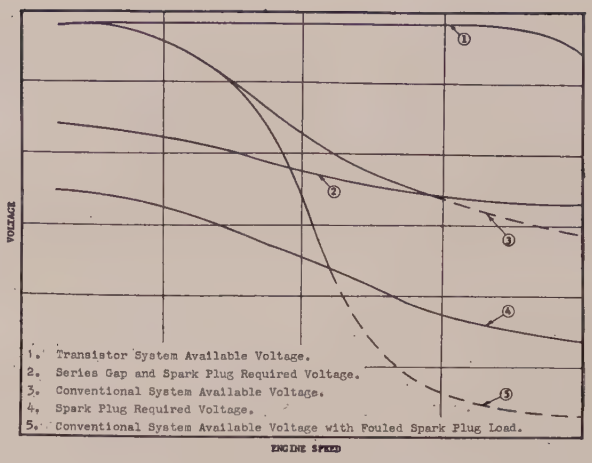

VOLTAGE
ENGINE SPEED
1. Transistor System Available Voltage.
2. Series Gap and Spark Plug Required Voltage.
3. Conventional System Available Voltage.
4. Spark Plug Required Voltage.
5. Conventional System Available Voltage with Fouled Spark Plug Load.

Fig. 17 — Typical available and required voltages for transistor switched, series gap, and conventional ignition systems over engine speed range

$$E = \tfrac{1}{2} Li^2 \qquad (1)$$

where: E = Energy, joules
L = Inductance, henries
i = Current, amp

From this formula, it follows logically that to maximize stored energy and thus coil output voltage, both the primary inductance and current should be made as large as possible.

This would be excellent, and coil design would be a straightforward art except for two factors — time and contact life. For reasonable contact life, the largest current that can be interrupted by conventional tungsten contacts has been empirically determined as about 5.0 amp at break. Currents higher than this will cause excessive arcing with resultant oxidation and/or large metal transfer.

The second factor, time, must be taken into account because, while the contact dwell period of a distributor may be considered a constant with respect to degrees of rotation, the contact dwell time is not. As engine speed increases, the dwell time becomes shorter; in fact, so short that the steady state relationships expressed in Ohm's Law no longer hold true. To make circuit calculations, we must modify Ohm's Law by adding the following transient factor $(1 - e^{-Rt/L})$, where:

e = Natural logarithm base
R = Primary circuit resistance, ohms
L = Primary inductance, henries
t = Time the distributor contacts are closed, sec

The combination is called Helmholtz Law:

$$i = \frac{E}{R}(1 - e^{-Rt/L}) \qquad (2)$$

Using the above relationship, it is clear that as engine speed increases and contact dwell time be-

comes shorter, primary current at break decreases. The larger the primary inductance, the greater the current decrease with speed. As coil performance depends on stored input energy and as the current in this relationship is a function of the second power as opposed to the inductance which is a function of the first power, it is obvious that it is more important to have a lower inductance and a high current at break than a high inductance and a dminishing current.

Nevertheless, the logic of using a primary inductance as high as possible is still cogent because of the limit in current at break imposed by contact life. The problem then splits into two distinct phases:

1. Secure a high primary inductance with a minimum current fall-off at high speeds.
2. Eliminate the current carrying restrictions of the contact set by devising a new concept of switching primary current.

APPENDIX II

The conventional ignition coil is basically an induction coil. It is, primarily, two oscillatory circuits loosely coupled together through a common magnetic circuit path. Loose coupling is employed so the highest possible voltage will be realized in the shortest possible time. By loose coupling is usually meant a long air gap.

In the induction coil as the contacts open a voltage is generated in the secondary by the relation:

$$e_2 = kL_2 \frac{di}{dt} \tag{3}$$

where:

k = Coefficient of coupling
L_2 = Secondary inductance
$\frac{di}{dt}$ = Rate of break of primary current

But also, there is a voltage developed across the primary:

$$e_1 = kL_1 \frac{di}{dt} \tag{4}$$

where: L_1 = Primary inductance

This primary induced voltage or "kickback" voltage tends to be high in conventional ignition coils. This inductive kickback is able to sustain an appreciably high current arc at the separating contacts resulting in "burning" and metal transfer.

With the "C" core type of coil, the magnetic path is almost closed but for a small air gap. There is a greater flux response and closer coupler between the primary and secondary, more energy is transferred from the primary ampere turns to build flux. Voltage output is proportional to the rate of change of flux and this can be shown in the companion equation for voltage:

$$e_2 = kN_2 \frac{d\phi}{dt} \tag{5}$$

where:

N_2 = Number of secondary turns
$\frac{d\phi}{dt}$ = Rate of collapse of flux

Since more energy is transferred to the secondary due to the close coupling, less energy is allowed to return to the primary by the collapsing flux. Consequently, the primary induced voltage is less. By proper choice of the design parameters, voltage can be kept to a safe limit.

APPENDIX III

Ignition designers have traditionally evaluated and compared coil designs by measuring "open circuit" available voltage and "open circuit" available voltage with a 1-megohm load.

The "open circuit" test is made with all spark gaps firing except the one being measured. In the laboratory, the distributed capacity of an automotive wiring harness is simulated by applying a capacitive load of approximately 50 μμf between the nonfiring or insulated lead and ground. The exact value of capacitive load applies may vary slightly among the various laboratories; but, in general, this is the type of test implied when an "open circuit" test is referred to.

In addition to "open circuit" data, the designer, traditionally, has also been interested in coil performance with a 1-megohm load paralleling the capacitive load on the nonfiring lead.

The "1-megohm" test is intended to simulate system performance with a fouled plug and has been an industry standard.

The objective of the induction coil designer has been to vary the various electrical parameters until he has optimized "open circuit" and "1-megohm" performance throughout the speed range.

References

1. "Notes on Sparking of Two Gaps in Series," by C. C. Paterson and N. R. Campbell. Great Britain Advisory Committee on Aeronautics, Internal Combustion Engine Subcommittee, Report 28, 1918.
2. "Subsidiary Gap as Means for Improving Ignition," by W. S. Gorton. NACA Report 57, 1919.
3. "Ignition Requirements for High-Compression Engines," by J. T. Fitzsimmons. SAE *Journal,* Vol. 24, March, 1929, pp. 306–314.
4. "Coil Ignition Systems," E. A. Watson. *Journal* of Institute of Electrical Engineers, (London), Vol. 70, 1932, p. 105.
5. "Analysis of High-Frequency Ignition Circuits," by A. W. Robinson, Jr. Electrical Engineer, Vol. 63, 1944, p. 916.
6. "Effect of Post-War Automotive Practices on Ignition Performance," by H. L. Hartzell. SAE Transactions, Vol. 53, 1945, pp. 427-431.
7. "Coming Car Engine Ignition," by H. L. Hartzell. SAE *Journal,* Vol. 55, July, 1947, p. 69.
8. "Electronic Ignition," by A. J. Poole. *Automotive Industries,* Vol. 96, Feb. 1, 1947, p. 37.
9. "Low Tension System Cures Ignition Ills," by A. C. Wall. SAE *Journal,* Vol. 55, January, 1947, p. 73.
10. "Ignition Systems for Future High-Compression Engines," by H. L. Hartzell. SAE *Journal,* Vol. 59, November, 1951, p. 45.
11. "Smitsvonk Low-Tension Capacity Ignition System," by W. Beye Smits and P. F. H. MacLane Pont. SAE *Journal,* Vol. 59, April, 1951, p. 61.
12. "Bibliography and Abstracts on Electrical Contacts, 1835–1951." ASTM Special Technical Publication 56-G, 1952.
13. "Expenditure of Current and Energy Required for Ignition in Explosion Engines," by C. C. Paterson and N. R. Campbell. Great Britain Advisory Commission on Aeronautics, Internal Combustion Engine Subcommittee, Report 25, 1918.
14. "Heat Energy of Various Ignition Sparks," by F. B. Silsbee, L. B. Loeb, and E. L. Fonseca. NACA Report 56,

1919.

15. "On Form of Temperature Wave Spreading from Point and Spherical Sources, with Suggested Application to Problem of Spark Ignition," by E. Taylor-Jones, J. D. Morgan, and R. V. Wheeler. *Philosophy Magazine,* Vol. 43, 1922, p. 359.

16. "Some Effects of Fuels and Lubricants on Auto-Ignition in Cars on Road," by R. K. Williams and J. R. Landis. SAE Transactions, Vol. 62, 1954, pp. 57–65.

17. "Correlation of Engines Noises with Combustion Phenomena," by R. Meagher, R. L. Johnson, and K. G. Parthemore. SAE Transactions, Vol. 63, 1955, pp. 481–492. (Abstract: SAE *Journal,* Vol. 63, March, 1955, p. 29.)

18. "On Rate of Rise of Secondary Potential of Ignition Systems," by G. E. Bairsto. Great Britain Advisory Committee on Aeronautics, Internal Combustion Engine Subcommittee, Report 31, 1918.

19. "On Effect of Shunted Resistance or Plug Leakage on Sparking Performance of Ignition System," by G. E. Bairsto. Great Britain Advisory Committee on Aeronautics, Internal Combustion Engine Subcommittee, Report 34, 1918.

20. "Benefits from Longer Gaps in Spark Plugs," by H. Rabezzana and H. Kalmar. *Automotive Industries,* Vol. 76, 1937, p. 222. *Automobile Engineer,* Vol. 27, 1937, p. 260.

21. "Effect of Internal Resistors on Spark Plug Electrode Erosion." Ethyl Corp. Report AR 101, 1943.

22. "Effect of Ignition Source on Spark Plug Electrode Erosion." Ethyl Corp. Report AR 102, 1943.

23. "Effect of Spark Polarity on Spark Plug Electrode Erosion." Ethyl Corp Report AR 103, 1943.

24. "Effect of Spark Polarity on Sparking Voltage during Actual Engine Operation." Ethyl Corp. Report AR 140, 1945.

25. "Effect of Varying Load Conditions on Sparking Voltage Requirements of Current Production Ceramic Spark Plugs." Ethyl Corp. Report AR 148, 1945.

26. "Modern Spark Plugs Permit Wider Gaps," by C. Cipriani. SAE *Journal,* Vol. 58, June, 1950, p. 65.

27. "Spark Plug Problems in Automotive Service," by R. C. Beaubier, H. J. Chalk, and M. M. Roensch. Paper presented at SAE National Passenger Car, Body, and Materials Meeting, Detroit, March, 1952.

28. "Spark Plug Fouling Studies," by R. J. Greenshields. SAE Transactions, Vol. 61, 1953, pp. 3–13.

29. "Spark Plug Fouling and Preignition." *Petroleum Processing,* Vol. 9, 1954, p. 53.

30. "Temperature Governs Spark Plug Performance." SAE *Journal,* Vol. 63, December, 1955, p. 107.

31. "Electric Contacts," by Ragnar Holm. Almqvist & Wiksells Akademiska Handbocker, 1946.

32. "Survey of Modern Ignition Systems." Electric Autolite Report.

33. "Theory of Induction Coil," 1923; and "Induction Coil," 1932, by E. Taylor Jones. Pub. by Sir Isaac Pitman and Sons, Ltd., London.

34. "Material Transfer in Opening Electrical Contacts," by Else Holm. Paper presented at Symposium on Electric Contacts, The Pennsylvania State University, November, 1953.

35. "Optimum Rate of Voltage Rise for Minimum Energy Loss in Ignition Systems," by M. F. Peters and L. H. Middleton. SAE *Quarterly* Transactions, Vol. 5, July, 1951, pp. 309–315.

DISCUSSION

New Contact Assembly Design Lengthens Life of Ignition System

— W. D. Walther
The Dayton Steel Foundry Co.

THE AUTHOR'S proposals for the various applications of transistorized ignition systems will be seriously considered in many types of engines in the years to come.

A quicker solution to the problems in the conventional ignition system may be obtained in an alternate manner. The author stated that, "The largest current that can be interrupted by the conventional tungsten contacts has been empirically determined at about 5 amp." For the standard ignition points, this means a current density in the order 500–2000 amp/sq in. of tungsten contact area.

By a unique change in design from a clapper-type ignition point to a valve actuating assembly the *effective* contact area transferring the energy has been increased at least tenfold; thereby, reducing current densities to 50–200 amp/sq in. In this new design of contact assembly the mechanical forces at the contact area are reduced in the same proportion. This permits (1) longer life of contacts and/or (2) a potential for more severe service. Fig. A shows such an assembly.

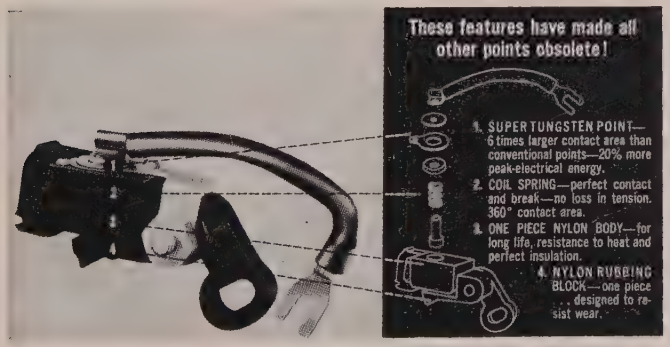

These features have made all other points obsolete!
1. SUPER TUNGSTEN POINT—6 times larger contact area than conventional points—20% more peak-electrical energy.
2. COIL SPRING—perfect contact and break—no loss in tension. 360° contact area.
3. ONE PIECE NYLON BODY—for long life, resistance to heat and perfect insulation.
4. NYLON RUBBING BLOCK—one piece . . . designed to resist wear.

Fig. A — Contact assembly

Such ignition contact assemblies (or switches) are now being produced. They have been service tested for the past four years in racing cars, trucks, and automobiles; and they exhibit outstanding performance. Such assemblies have been proved to give higher voltage, no point bounce, up to 8500 engine rpm, and long life.

ORAL DISCUSSION

Reported by D. W. Morrison
Interstate Commerce Commission

Orval Brouer, Swift & Co.: What is your opinion of the "rim-fire" plugs being advertised?

Mr. Spaulding: These plugs are usually called "fuel igniter" plugs. They do not have a conventional spark-plug gap configuration. The spark in this type plug jumps an air gap from the center electrode to the shell; there is no side electrode attached to the shell.

The manufacturer's advertising implies that the surface discharge principle of low voltage ignition systems is used; you will note, however, that there is an air gap that the spark is required to jump. They are used in connection with conventional ignition systems. They are not "surface discharge" or low-voltage spark plugs as defined in my paper.

F. W. Petring, Bureau of Public Roads: What economic conditions would be used in determining whether transistorized ignition systems should be used—fleet operation or high reliability?

Mr. Spaulding: Generally, neither of the two by itself. There are a number of things to be taken into consideration: down time, initial cost, bookkeeping accounting. An analysis of the operation would have to be made before a decision could be made.

Correcting Horsepower

a realistic method

A. K. Blackwood, International Harvester Co.
W. J. McCulla, Caterpillar Tractor Co.

THE SAE DIESEL Engine Test Code adopted in January, 1931, and last revised in January, 1941, has not been accepted as a standard by the diesel industry. To consider possible revisions, the SAE Diesel Engine Test Code Committee was organized as a subcommittee of the SAE Engine Committee on Nov. 22, 1955. This subcommittee established as the code's general purpose: "to make available standards covering dynamometer laboratory tests for determining and presenting performance characteristics of diesel engines which are designed primarily for industrial and transportation applications."

Specifically, the code was to be designed to accomplish the following:

1. Establish standardized testing procedures, methods for reporting results, and specifications of equipment and instrumentation.
2. Develop a consistent set of definitions to describe the terms used in connection with engine tests and power output.
3. Clarify the concepts of correcting the power output of diesel engines for variations in ambient test conditions, and develop sound engineering methods for accomplishing this correction.

A progress report was presented in June, 1958. This paper[1] covered items 1 and 2 in detail. It is the purpose of this report to describe the correction method, modes of application, and limitations, and to furnish actual engine test data to support its adoption.

Need for Correcting Power Output

Operating data show that the power output of a diesel engine can be affected by variations in the charge density caused by changes in barometric pressure, temperature, and humidity of the ambient atmosphere. Therefore, it is necessary to have a means for converting measurements of power output at atmospheric conditions existing during the test to the power output that the engine would deliver under standard atmospheric conditions.

Given a uniform basis for comparing engine output, the engine user can select an engine of the power to match his application and the engine manufacturer can use this standard to maintain the quality of his product.

Correction methods are required to fit two different conditions of operation. We refer to these as the constant air/fuel ratio method and the constant fuel rate method. The constant air/fuel ratio method is a means of determining the power output at standard ambient air conditions when the fuel input to the engine is varied in proportion to the air density. In this method, then, the air/fuel ratio at test and standard conditions is constant. It is used to determine the engine power output at standard ambient conditions at the same smoke level as at the test conditions. As an example, in the development of an engine, constant speed-variable load curves are often used for comparison of engine performance. To eliminate the effects of day-to-day changes in ambient conditions, these curves may be corrected by the constant air/fuel ratio method and directly compared at the standard conditions for the effect of the design change under consideration.

The constant fuel rate method is used to determine the engine power output at standard ambient air conditions when the fuel input to the engine is fixed at some predetermined value. With constant fuel input, the air/fuel ratio and the smoke level will differ at test and standard conditions. In a dragline application, for example, the fuel rate at each speed is set at a limited quantity by a fuel stop. It may be necessary to set the engine for this use under ambient conditions other than standard. The constant fuel rate method may then be used to determine the horsepower setting at each speed

---

* Paper presented at SAE Annual Meeting, Detroit, Jan. 12, 1960.
[1] "Development of Test Code and Correction Factors for Diesel Engines — A Progress Report," by M. A. Elliott and A. K. Blackwood. Paper presented at SAE Summer Meeting, Atlantic City, June 12, 1958.

Output

for diesel engines

THIS PAPER outlines the corrections to the SAE Diesel Engine Test Code approved by the SAE Technical Board.

The basic correction method is the constant air/fuel ratio method. This method is intended to determine the output under standard conditions when the fuel rate is adjusted to maintain equivalent combustion conditions or exhaust smoke density.

The method is based on the assumptions that the following exist at test and standard conditions: constant volumetric efficiency, constant combustion efficiency at a given air/fuel ratio, same friction horsepower at test and standard conditions.*

under the test conditions which will result in the specified horsepower at standard conditions.

Standard Conditions

The accuracy of any correction method is limited by the validity and universality of the assumptions made in its development. Inaccuracies are minimized by maintaining the required correction at low numerical values. An obvious way to approach this is to select standard ambient conditions as close as possible to those under which the tests will be conducted. To that end a study was made of the ambient conditions existing in the testing facilities of various diesel-engine manufacturers, of conditions existing across the United States in areas of maximum engine use, and of conditions being used in other areas. Standard ambient conditions in use in various laboratories were surprisingly similar. An analysis of the ambient conditions existing in a typical laboratory is shown in Figs. 1 and 2. Fig. 1 shows that the average barometric pressure over a year's time was 29.25 in. of Hg, and that 90% of all the measurements of barometric pressures were within ±0.4 in. of Hg of the average. Fig. 1 also shows that the average temperature over a year's time was 83 F and that 90% of all measurements were within ±10 F of the average. Fig. 2 shows the average vapor pressure over a year's time to be 0.371 in. of Hg, 75% of all measurements being within ±0.2 in. of Hg. Based on the average ambient conditions, an average dry air density of 0.0702 lb cu ft existed in this particular laboratory. The present "unofficial" standard barometer and temperature are also shown on these curves. It will be noted that a pressure of 29.92 in. of Hg or above occurred less than 1% of the time and that the temperature never approached 60 F in this particular laboratory.

A study of the population distribution in the United States in relation to the dry air density during the months of June through September showed that 84% of the population lived in areas in which the dry air density was between 0.068 and 0.072 lb cu ft. This same study showed the distribution of trucks and tractors in the United States to be approximately proportional to the population distribution.

On the basis of these studies and the general philosophy of keeping test conditions close to standard conditions, the committee agreed that the standard conditions should correspond to an altitude of 500 ft (barometric pressure of 29.38 in. of Hg), a temperature of 85 F and a relative humidity of 50%. This results in a dry barometer of 28.77 in. of Hg, which with the temperature of 85 F corresponds to a dry air density of 0.070 lb cu ft. These conditions

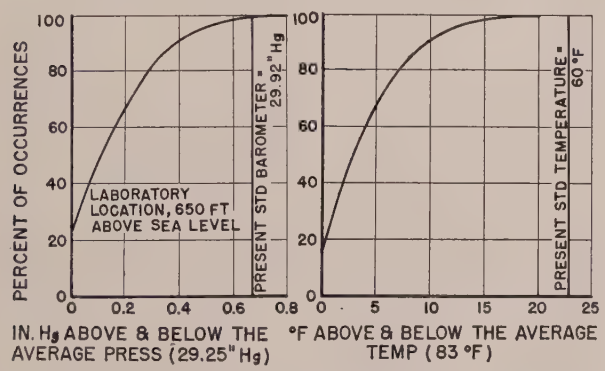

PERCENT OF OCCURRENCES
100
80
60
40
20
0
LABORATORY LOCATION, 650 FT ABOVE SEA LEVEL
PRESENT STD BAROMETER = 29.92" Hg
0 0.2 0.4 0.6 0.8
IN. Hg ABOVE & BELOW THE AVERAGE PRESS (29.25" Hg)
PRESENT STD TEMPERATURE = 60°F
0 5 10 15 20 25
°F ABOVE & BELOW THE AVERAGE TEMP (83 °F)

Fig. 1 — Barometric pressure (left) and ambient temperature (right), per cent of occurrences in relation to deviation from average

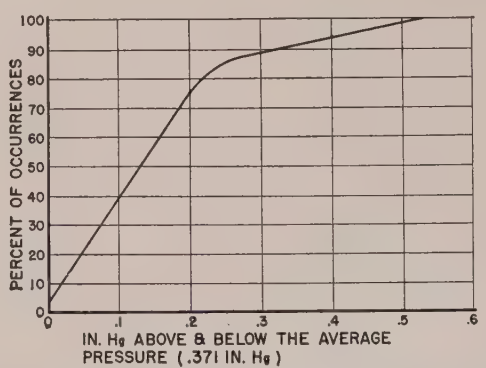


Fig. 2 — Vapor pressure, per cent of occurrences in relation to deviation from average

are very close to those specified in British Standard 649:1958.

Present SAE Correction Method

As shown in the Diesel Engine Test Code, the presently approved formula for correcting engine output for variations in ambient conditions is:

$$\text{bhp}_C = \text{bhp}_O \times \frac{30.212}{P_O} \times \sqrt{\frac{T_O}{550}} \qquad (1)$$

where:

bhp_C = Corrected brake horsepower
bhp_O = Observed brake horsepower
P_O = Observed barometric pressure, in. of Hg
T_O = Observed absolute temperature, F

This correction method, which purports to correct output to standard conditions of 30.212 in. of Hg[2] and 90 F, is widely used only in modified form, to a base of 29.92 in. of Hg and 60 F in this form:

$$\text{bhp}_C = \text{bhp}_O \times \frac{29.92}{P_O} \times \sqrt{\frac{T_O}{520}} \qquad (2)$$

This formula, when applied to carbureted engines, is approximately correct throughout the load range. Its inaccuracy in correcting diesel-engine output is related to the widely varying air/fuel ratio of the diesel.

Consider the familiar part load fuel consumption curve of a diesel engine (Fig. 3). At a constant speed the airflow is practically the same from no load to full load, while fuel input varies possibly 5 or 6 to 1. About point *A*, for instance, specific fuel consumption is essentially constant over a considerable load range. This means that increasing or decreasing the flow of fuel brings about a proportionate change in horsepower. Enough excess air is present in the cylinder so that each droplet of fuel finds the oxygen required for burning. If the fuel rate is held constant, it makes little difference in output if the available air is varied 20% or more. At maximum output, point *B*, an increase in fuel flow can produce no additional power, for all the air is being used to the best of the engine's ability. A slight decrease in fuel flow would decrease power in only a fractional proportion. Under these conditions a change in available air would have a directly proportionate effect upon output. Simply restated, at part load the fuel rate controls the output; at maximum load the air supply controls the output. The presently used SAE correction method improperly provides a constant percentage correction to these widely varying conditions.

Other weaknesses in the current method are:

1. It does not distinguish between correction at constant fuel rate and at constant fuel/air ratio.
2. It does not recognize differences between engines in their ability to make use of available air.
3. The approved SAE formula assumes 50% relative humidity at test conditions, dry air at standard conditions. The modified formula ignores humidity.
4. Base conditions in both the SAE and the modified formula differ appreciably from average laboratory conditions, causing correction errors to be greater than otherwise necessary.
5. The modified formula is widely used only as an expedient because no better method has been given official SAE sanction. Because it is known to many people, it appears in the sales literature of various manufacturers and is often used on their production test floors. These same manufacturers usually use a more accurate method in their laboratory work.

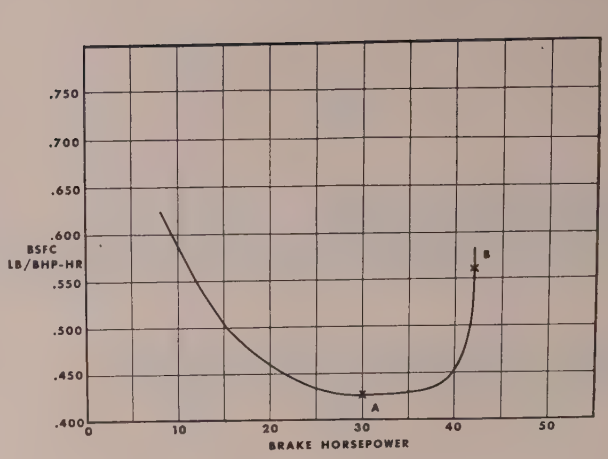


Fig. 3 — Part-load fuel consumption

Proposed SAE Correction Method

With these limitations of the present system in mind the subcommittee set about developing a method that would provide reasonable accuracy and yet not require complicated formulas or elaborate instrumentation.

The basic correction method proposed is the constant air/fuel ratio method. This method is intended to determine the output under standard conditions when the fuel rate is adjusted to maintain equivalent combustion conditions or exhaust smoke density. In developing this method it is assumed that certain relationships exist at test and standard conditions.

First, *the volumetric efficiency is assumed constant.* This means that the weight of air available for combustion is directly proportional to ambient air density. Since air/fuel ratio is held constant, the weight of fuel put into the engine is also proportional to the ambient air density.

Combustion efficiency at a given air/fuel ratio is assumed constant. Therefore, the indicated horsepower is directly proportional to the weight of fuel introduced. It follows, then, that at constant air/fuel ratio the indicated horsepower is directly proportional to the ambient air density.

This leads us to the relationship:

$$\text{ihp}_s = \text{ihp}_t \frac{d_s}{d_t} \qquad (3)$$

where:

ihp_s = Indicated horsepower at standard conditions
ihp_t = Indicated horsepower at test conditions
d_s = Density of dry air at standard conditions
d_t = Density of dry air at test conditions

Since what we are looking for is brake rather than indicated horsepower, the further assumption is made that *friction horsepower is the same at test*

[2] The figure 30.212 includes a humidity factor intended to give a nominal correction from 50% relative humidity at test conditions to dry air at sea level

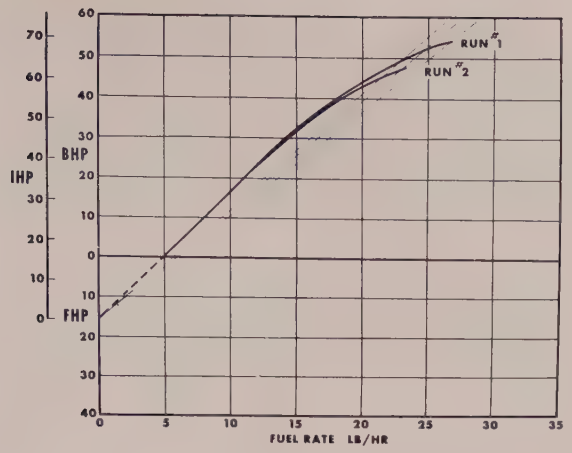


Fig. 4 — Brake horsepower versus fuel rate

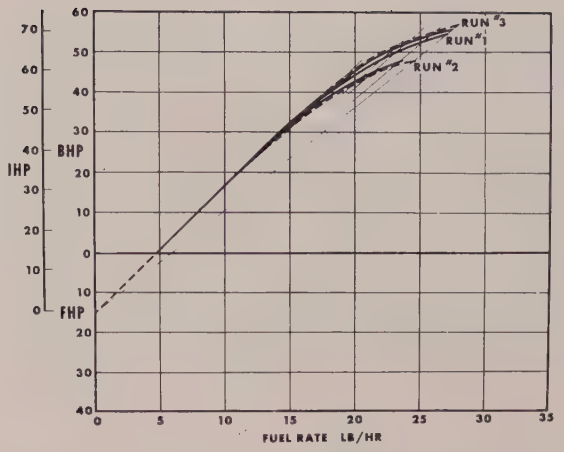


Fig. 5 — Brake horsepower versus fuel rate

and standard conditions. The relationship then is:

$$\text{bhp}_s = (\text{bhp}_t + \text{fhp}_t)\frac{d_s}{d_t} - \text{fhp}_t \qquad (4)$$

where:

bhp_s = Brake horsepower at standard conditions
bhp_t = Brake horsepower at test conditions
fhp_t = Friction horsepower at test conditions

The density ratio $\frac{d_s}{d_t}$ of the ambient air at standard and test conditions is equal to the product of the dry barometer pressure ratio $\frac{P_{sd}}{P_{td}}$ and the absolute temperature ratio $\frac{T_t}{T_s}$, or:

$$\frac{d_s}{d_t} = \frac{P_{sd}}{P_{td}} \times \frac{T_t}{T_s} \qquad (5)$$

If d_s is shown as 0.070 lb/cu ft then d_t is calculated and expressed in the same units:

$$d_t = 1.326 \times \frac{P_{td}}{T_t} \qquad (6)$$

Note that what really interests the engine is the oxygen ratio. If the density ratio is truly to represent this, the observed barometer reading must be reduced by the partial pressure of water vapor in the air.

This being a constant air/fuel ratio correction, the fuel rate must vary directly as the density. Thus:

$$F_s = F_t \frac{d_s}{d_t} \qquad (7)$$

where:

F_s = Fuel rate lb/hr at standard conditions
F_t = Fuel rate lb/hr at test conditions

If correction on a brake specific fuel consumption basis is preferred, the fuel consumption under standard conditions can be calculated, after bhp_s is found, by the related formula:

$$\text{bsfc}_s = \text{bsfc}_t \times \frac{\text{bhp}_t}{\text{bhp}_s} \times \frac{d_s}{d_t} \qquad (8)$$

These relationships can be shown most plainly

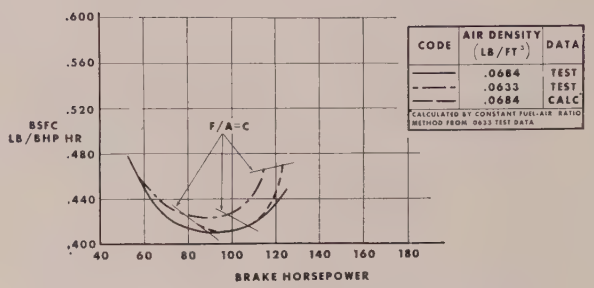


Fig. 6 — Two-cycle diesel engine, comparison of test and calculated data at 1600 rpm

by referring to a plot of brake horsepower versus fuel rate such as is used to determine friction horsepower by extrapolation. Fig. 4 is a plot of actual data taken on a four cycle engine at constant speed. Ambient conditions were:

	Run No. 1	Run No. 2
Observed barometer, in. of Hg	30.04	29.73
Temperature, F	120	119
Relative humidity %	11	85

The light diagonal lines radiating from the origin of the ihp — fuel rate scales are lines of constant air/fuel ratio. If all our assumptions are perfectly valid, the points of intersection of a constant air/fuel ratio line with the two fuel rate curves will have the relationship indicated by the proposed correction method. Points on the curve of Run 2 could thus have been calculated from points on Run 1. Similarly, bhp at constant fuel rate can be read at the points of intersection of any vertical line with the two curves.

In Fig. 5 are shown results of a third run on the engine of Fig. 4. Ambient conditions were 29.53 in. of Hg, 90 F, and 22.5% relative humidity. Dotted lines are calculated values of Runs 2 and 3 based upon Run 1 data. Air density differences between Runs 1 and 2 were caused almost entirely by humidity; between Runs 1 and 3 temperature was the major factor.

If preferred, data can be plotted on a specific fuel consumption basis. Figs. 6–8 present data of a two cycle engine in this way. In these curves test data

run at an air density of 0.0633 lb/ft³ are corrected to 0.0684 lb/ft³ and compared to actual test data at that density. To determine output at constant fuel rate on this plot, it is necessary to construct constant fuel rate lines, as in Fig. 7. Note that these lines are not straight, but are somewhat concave upward.

Results of a simulated altitude test of a 4-stroke engine are shown in Fig. 9. These curves were run under ambient humidity conditions, which were essentially constant. Temperature was controlled at 90 F. The air inlet was restricted and the exhaust evacuated to the dry barometer values shown. Although there is obvious disagreement between calculated and observed outputs, it will be noted that the runs compared differed in pressure by 3⅓ and 4⅓ in. of Hg.

Runs at three engine speeds under ambient conditions of 0.0675 and 0.0715 lb/ft³ are shown in Fig.

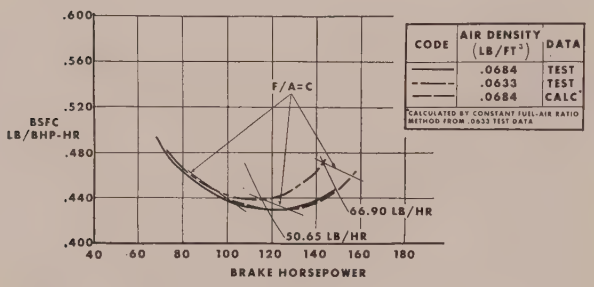


Fig. 7 — Two-cycle diesel engine, comparison of test and calculated data at 2100 rpm

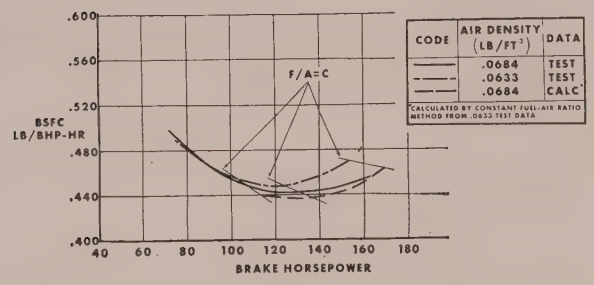


Fig. 8 — Two-cycle diesel engine, comparison of test and calculated data at 2300 rpm

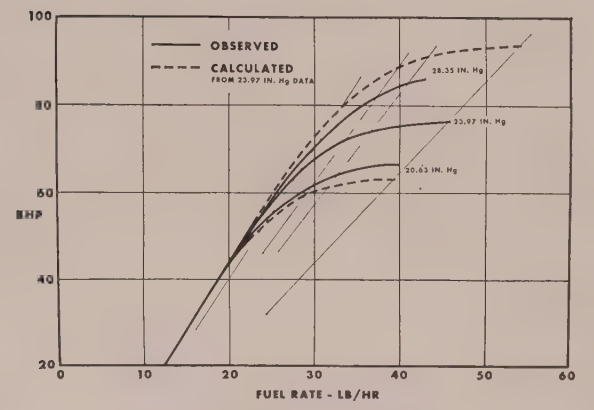


Fig. 9 — Four-cycle diesel engine

10. The engine was an aircooled 4-stroke diesel. Calculated outputs agreed rather well with observed horsepowers.

Data for Fig. 11 were taken from a paper by Martin A. Elliott.[3] Although temperature was a variable, pressure accounted for most of the difference in density. This 4-stroke swirl chamber diesel showed excellent correlation of test and calculated data.

Application of Proposed Correction Method

An effort was made to select the preceding examples so that densities were far enough apart to minimize the effect of experimental errors, yet not so different that our basic assumptions were invalidated. In each case, one test condition was assumed standard and a complete part-load curve "corrected" from another condition for the sake of comparison.

Normal use of the method is much simpler. Where only a point requires correction under constant air/fuel ratio conditions and where friction horsepower is known, no plotting is required. A form of the equation including standard conditions should then be used:

$$\text{bhp}_s = (\text{bhp}_t + \text{fhp}_t)\left(\frac{28.77}{P_{td}}\right)\left(\frac{t_t + 460}{545}\right) - \text{fhp}_t \qquad (9)$$

and:

$$F_s = F_t\left(\frac{28.77}{P_{td}}\right)\left(\frac{t_t + 460}{545}\right) \qquad (10)$$

or:

$$\text{bsfc}_s = \text{bsfc}_t \times \frac{\text{bhp}_t}{\text{bhp}_s}\left(\frac{28.77}{P_{td}}\right)\left(\frac{t_t + 460}{545}\right) \qquad (11)$$

Corrected output at constant fuel delivery is determined graphically from constant air/fuel ratio calculations. An accurate determination requires that data be taken at no less than three points. These three points will establish the approximate curvature of the fuel consumption line near the required output.

One of the test points should be at the desired horsepower. When the test density is below the standard density, the other two points should be at approximately 5 and 10% lower fuel rates. When the test density is above the standard density, the other two points for greatest accuracy should be at approximately 5 and 10% higher fuel rates. When practical considerations such as limited pump delivery prevent use of a higher fuel rate, points below the fuel rate at the desired horsepower may be used. Some accuracy is then sacrificed because extrapolation, rather than interpolation, is involved.

The brake horsepower at each of the three points is corrected to standard conditions and the corresponding fuel rates are calculated by use of two of Eqs. 9–11 (constant air/fuel ratio method).

A graph is constructed with the brake horsepower as ordinate and fuel rate as abscissa. A smooth curve is drawn through the three points at test ambient conditions and a similar curve through the three points at standard ambient conditions. The brake horsepower at standard conditions is then read from the graph at the intersection of the

[3] "Conversion of Measurements of Power Output of Diesel Engines to Standard Atmosphere Conditions," by Martin A. Elliott, ASME Transactions, July, 1946, p. 525.

test fuel rate and the curve for the standard density.

Alternatively, the graph is constructed with the brake horsepower as abscissa and the brake specific fuel consumption as ordinate. Smooth curves are drawn through the three test points and through the three corrected points. A constant fuel rate line is constructed through the point to be corrected. Points on this line are a constant product of bhp and bsfc. Intersection of this line with the corrected curve gives the corrected horsepower.

The following example of a constant fuel delivery correction is adapted from the proposal submitted to the SAE Engine Committee for approval.

Required: Output at standard conditions at the same fuel rate that produces 82.0 bhp at test conditions. Ambient density less than standard density.

Test Points: Three test points required, at 82.0 bhp and at fuel rates approximately 5 and 10% lower than the fuel rate at 82.0 bhp.

Observed Data:

	bhp_t	F_t (lb/hr)
Test Point 1	82.00	38.35
Test Point 2	78.25	36.43
Test Point 3	74.10	34.51

$fhp_t = 32.2$
Barometer at test conditions = 28.50 in. of Hg
Dry bulb temperature at test conditions = 95 F
Wet bulb temperature at test conditions = 60 F

Calculations —

Vapor pressure from psychometric chart 0.15 in. of Hg

Dry barometer at test conditions

$$28.50 - 0.15 = 28.35$$

At Point 1:

$$bhp_s = (bhp_t + fhp_t)\left(\frac{28.77}{P_{td}}\right)\left(\frac{t_t + 460}{545}\right) - fhp_t \qquad (9)$$

$$= (82.00 + 32.20)\left(\frac{28.77}{28.35}\right)\left(\frac{95 + 460}{545}\right) - 32.2$$

$$= 85.8 \text{ bhp}$$

$$F_s = F_t\left(\frac{28.77}{P_{td}}\right)\left(\frac{t_t + 460}{545}\right) \qquad (10)$$

$$= 38.35\left(\frac{28.77}{28.35}\right)\left(\frac{95 + 460}{545}\right)$$

$$= 39.6 \text{ lb/hr}$$

Points 2 and 3 are also corrected by the constant air/fuel ratio method.

	bph_s	F_s (lb/hr)
Std. Point 1	85.8	39.6
Std. Point 2	82.0	37.65
Std. Point 3	77.7	35.65

The brake horsepower and fuel rate for the three test and corrected points are plotted as shown in Fig. 12. Since bhp_t at Test point 1 is the horsepower to be corrected, the brake horsepower at standard conditions and constant fuel delivery is read from the graph at the intersection of the constant fuel rate line through Test Point 1 and the curve for standard conditions.

Corrected bhp from graph = 83.4
Fuel rate, same as at Test Point 1 = 38.35 lb/hr

$$\text{Specific fuel consumption} = \frac{F}{bhp} = \frac{38.35}{83.4}$$

$$= 0.460 \text{ lb/bhp-hr}$$

Simplified Production Method — In general, the correction method used on the production test floor, where engines are being adjusted for specific applications, will be the constant fuel rate method. As previously described, it would appear to involve more work than could be justified on a production basis. Various means of simplifying the approach, by making up charts based on a sampling of the particular model engine under test, are possible with very little sacrifice in accuracy being involved.

One such method is described. Part-load fuel consumption curves are run on a number of production engines at a selected speed. These curves are

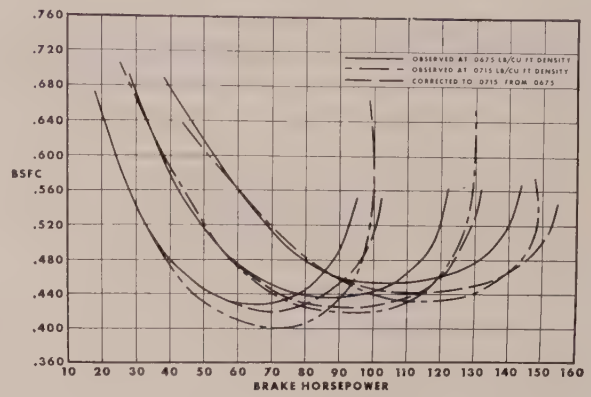


Fig. 10 — Aircooled 4-cycle diesel engine, part-load fuel consumption at three speeds

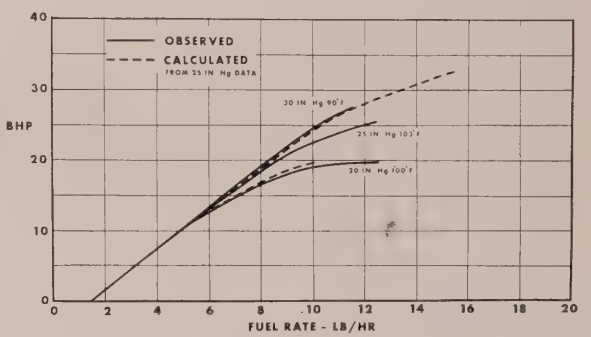


Fig. 11 — Engine *B* (from Elliott)

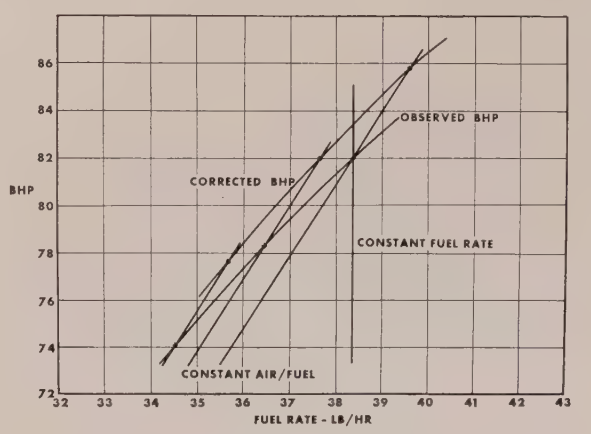


Fig. 12 — Constant fuel rate correction

corrected to a common density by the constant air/fuel ratio method and an average curve selected. This curve is then corrected by the same method to a series of densities, including standard density, covering the range normally encountered on the test floor. If the horsepower or dynamometer beam load required for a particular application is spotted on the standard density line and a constant fuel rate line is drawn through this point, we then have a line which gives us output versus density at a constant fuel rate. (See Fig. 13.) A graph of these data can then be plotted in many ways. A handy method for selecting the required bhp or beam load at any given dry barometer and ambient temperature is shown in Fig. 14. This plots beam load as ordinate and dry barometer as abscissa. Constant temperature lines are plotted at such an angle that the intersection of abscissa and temperature line on any horizontal line gives a constant dry air density. This plot can be applied with standard production fuel and bhp tolerances.

Limitations — At present, the correction method is applicable to naturally aspirated and mechanically blown diesel engines but not to turbocharged engines.

Two of the assumptions upon which this correction method is based are that at constant air/fuel ratio the combustion efficiency and volumetric efficiency are constant. This will be true provided the range of test pressures and temperatures is small enough so that other engine operating characteristics such as ignition lag, rate of combustion, peak pressure, heat rejection, and the like, are not adversely affected. The test conditions at which this influence becomes apparent cannot be accurately defined, as they will depend upon fuel characteristics, engine size, speed range, combustion-chamber design, and the rest. Any error introduced by the assumptions will, therefore, increase as the spread between the test and standard conditions increases. For this reason it is recommended that the test data be obtained with barometric pressures between 28 and 31 in. of Hg, and with inlet air temperature between 40 and 120 F.

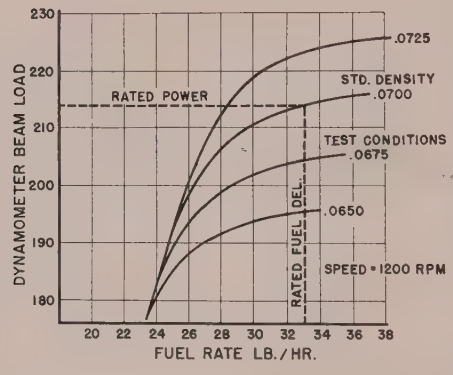


Fig. 13 — Constant speed variable load curves for simplified constant fuel rate correction method

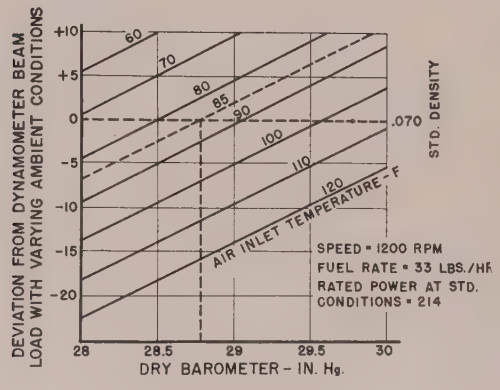


Fig. 14 — Simplified constant fuel rate correction method applicable to specific production engine

Conclusions

It may be concluded from the data presented herein that the correction method being recommended by this committee, when used within the limitations set forth, fulfills the requirement for an accurate method based on sound engineering practice for determining the power output of an engine at standard ambient conditions when tested at ambient conditions other than standard. It may be used either where the air/fuel ratio is maintained constant or where the fuel rate is maintained constant and is readily adaptable through procedures described to present production test requirements.

This study is being expanded to include turbocharged engines, in which the effects of variations of ambient conditions not only on the engine must be taken into account, but also their effects on the turbocharger and on the turbochanger-engine match. To this end a new subcommittee is being formed, including as a nucleus the members of the present subcommittee who have experience in turbocharging plus additional members who can aid in the solution of this problem. The correction method described here, modified by a factor taking into account the characteristics of turbochargers under varying ambient conditions, is a starting point for the continued study.

The Diesel Engine Test Code (including the correction method described here) as developed by the Diesel Engine Test Code Subcommittee has been approved by the Engine Committee and the SAE Technical Board.

Acknowledgments

The authors are indebted to the members of the Diesel Engine Test Code Subcommittee who participated in the development of this correction method and the selection of standard conditions and who submitted test data in support of the study. Membership of this committee, in addition to the authors, includes: M. A. Elliott, Chairman, Inst. of Gas Technology, Chicago, Ill.; W. F. Burrows, White Motor Co., Diesel Engine Div., Springfield, Ohio; H. A. Cook, Thompson Ramo Wooldridge, Inc., Cleveland, Ohio; W. A. Engstrom, Continental Motors Corp., Muskegon, Mich.; M. L. Fast, Cummins Engine Co., Inc., Columbus, Ind.; W. I. Hamilton, Allis-Chalmers Mfg. Co., Harvey, Ill.; S. Kalmar, General Motors Corp., Engrg. Staff, Detroit, Mich.; L. F. Larsen, University of Nebraska, Lincoln, Nebr.; S. E. Lytle, U. S. Army, Corps of Engineers, Ft. Belvoir, Va.; D. E. Miller, Detroit Diesel Engine Div., GMC, Detroit, Mich.; H. M. Wiles, Waukesha Motor Co., Waukesha, Wis.; A. H. Zeitz, Jr., Ethyl Corp., Detroit, Mich.

properties of Asbestos Reinforced Laminates at elevated temperatures

Norman E. Wahl
Cornell Aeronautical Laboratory, Inc.

IF ROCKET OR MISSILE designers were asked to choose one specific property of engineering materials they would like to have improved, the largest percentage would undoubtedly select strength at high temperature.

In addition to retaining strength at high temperatures, missile materials must be resistant to erosion and ablation. Missile structures must also be satisfactory when subjected to aerodynamic and acceleration loads, high stresses of vibration, and thermal shock. The need for low-density, easily fabricated, heat-resistant materials has resulted in a continuing search for more effective combinations of known materials, as well as the development of new materials.

This paper discusses some interesting results obtained in studies of composite materials that might be used for rocket or missile construction. The studies were conducted at Cornell Aeronautical Laboratory under Contract AF33(616)-2926, DA-30-115-ORD-543, AF33(616)-5683, and at Forest Products Laboratory under Contract DO-33(616)-58-1.*

THE TARGET OBJECTIVE in the study of composite materials was the development of a high-strength low-density material for use in such items as missile bodies, nose cones, or combustion chambers of rocket motors. The service conditions to which these components might be exposed varied considerably.

It was estimated that missile bodies might attain a maximum temperature in the range of 500–1000 F for relatively long periods of time, while the nose cones and combustion chambers might reach 5000 F or higher for a few seconds.

The aerodynamic heating and ablation of the nose cone occurs only during high-speed flight through the earth's atmosphere, generally for a very short time. Components of liquid or solid-propellant rocket motors have to withstand generally more severe erosion conditions and higher temperatures. The heating conditions in rocket combustion chambers vary; since at the same gas temperature, the rate of heat transfer in burning a solid propellant is considerably greater than that in a liquid fuel motor, because of the radiation from incandescent solid particles. Thus, from the standpoint of temperature alone, conditions are very severe in solid-propellant motors. For example, ablation of rocket nozzles due to erosion can be more severe in liquid-propellant than in solid-propellant motors, because the exhaust gases are likely to be oxidizing.

In order to roughly simulate these service conditions in the evaluation of the composite materials under study, three types of tests were used:

1. Short-time exposure to rocket engine exhaust with varying percentages of oxygen.
2. Short-time exposure to ablation conditions with varying heat fluxes.
3. Long-time exposure to temperatures up to 1000 F.

Composite Materials

Because preliminary studies at Cornell Aeronautical Laboratory had indicated that combinations of long fibered asbestos and other high-temperature resistant fillers for reinforcement of phenolic or inorganic silicone resins resulted in a structural plastic material that had a high strength-to-weight ratio,

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 8, 1959.

good-to-excellent resistance to erosion and ablation, and maintained most of its strength at temperatures as high as 4000 F for exposure periods up to 30 sec, this study of asbestos composites was undertaken.

Phenolic-Asbestos Laminates — A group of ⅛- and ½-in. thick laminate panels of phenolic preimpregnated asbestos felt were molded in a press at pressures of 15, 30, 60, 200, and 400 psi for one-half hour at 325 F.

The panels molded at 15 and 30 psi were relatively thick, appeared dry, and had little rigidity. These characteristics were attributed to insufficient pressure to cause the phenolic asbestos plies to fuse to a homogeneous mass. It should be recognized that at low pressures, the preimpregnated asbestos sheet should have the highest percentage of resin flow obtainable. If a preimpregnated asbestos felt with a higher flow had been used at molding pressures of 15 and 30 psi, the panels might have fused and would undoubtedly have better physical properties.

After the panels were molded at 325 F for one half hour, they were postcured in an air-circulating oven for 4 hr at 300 and 350 F, followed by 16 hr at 400 F before evaluation.

An index of the optimum molding pressure, is the water absorption after 24-hr immersion. This property was determined on the first group of ⅛-in. laminates molded. Table 1 indicates that the asbestos phenolic laminates molded at 200- and 400-psi pressure resulted in nonporous dense laminates.

Table 1 — Water Absorption of Molded Asbestos Phenolic Laminates

Specimen No.	Molding Pressure, psi	Water Absorption, %
535	15	46.3
536	30	40.2
537	60	21.3
538	200	1.6
539	400	1.3

Table 2 — Water Absorption of Silicone Asbestos Laminates

Specimen No.	Material	Molding Pressure, psi	Water Absorption, %
503	9526 RB-5, DC-2106	30	4.6
513	9526 RB-5, DC-2106	60	3.4
518	9526 RB-5, DC-2106	200	2.2

Table 3 — Gas Temperatures

Condition of Test	I	II	III
Combustion Temperature R	4820	4060	3260
Excess Oxygen, %	5	15	20
Calculated Exhaust Gas Velocity		Mach 2.6	

Subsequent test panels were all molded at 400 psi.

Silicone Asbestos Laminates — In the preliminary work, silicone asbestos laminates were made by immersing sheets of 18 in. × 19 in. 40-RPD asbestos in catalyzed DC-2106 silicone resin from Dow Corning Corp. The solvent, which was toluene, was removed from the sheets by heating in an oven at 225 F for 15 min. This operation was very cumbersome because of the extremely fragile nature of the asbestos, which tore and stuck to anything the resin impregnated sheet would touch.

Panels molded of this impregnated 40-RPD asbestos molded in a press at 350 F for one hour appeared to be satisfactory, but when postcured 24 hr at 500 F blistered badly.

In further work with asbestos silicone laminates it was found that predrying of the impregnated sheet at 225 F for 3 min before molding at 350 F tended to prevent blistering of the laminate when postcured at 500 F for 24 hr.

Because of the difficulty of hand impregnation of asbestos sheet with silicone resin, all the panels fabricated in the last phase of the program were molded from silicone-preimpregnated asbestos sheet purchased from U. S. Polymeric Chemicals Inc.

Two panels of 40-RPD impregnated with DC-2106 silicone resin molded at 15 psi which were incompletely fused due to the low resin flow and pressure, were not tested. Two panels molded at 30 and 60 psi, had a dry appearance. The water absorption, after 24-hr immersion, of silicone asbestos laminates also tends to indicate that higher pressure gives denser and better laminates. (Table 2).

Based upon the results obtained in this study, it would appear that molding pressures below 200 psi for asbestos silicone laminates are not advisable. The optimum mechanical properties were obtained

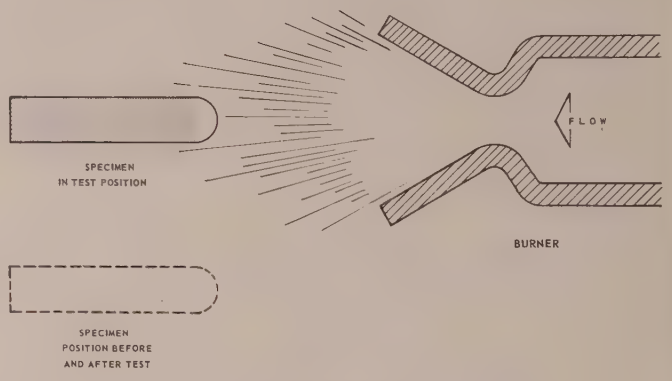


Fig. 1 — Profile view of specimen insertion in rocket exhaust

Table 4 — Asbestos-Reinforced Laminates — ½ in. Thick Exposed to Rocket Exhaust

Type Resin & Specimen No.	Leading Edge Radius, in.	Test Condition	Time Exposure, sec	Loss Weight grams/sec	Remarks
Phenolic 26	9/32	I	20.8	1.9	Leading edge eroded irregularly back ½ in. Sides ablated
Phenolic 27	9/32	II	28.0	1.7	Leading edge eroded back ¼ in. Sides ablated
Phenolic 28	9/32	III	20.3	1.4	Leading edge eroded back ¼ in. Sides ablated
Silicone 73	1/4	I	6.6	7.6	Specimens badly ablated showing tendency to delaminate
Silicone 74	1/4	II	6.5	5.0	"
Silicone 75	1/4	III	5.8	6.1	"
Silicone 76	5/16	II	3.5	7.6	Specimen showed almost no pyrolysis — considerable delamination of outer plies

if molding pressures of 200 psi or above were used. All silicone laminate panels of ⅛ or ½ in. thickness used in thermal evaluation tests were molded at 200-psi pressure at 350 F, then postcured at 500 F for 24 hr.

Tests and Results

Short-Time Exposure to High Temperature with Varying Percentages of Oxygen — The source of the very high-temperature, high-velocity gas for simulating rocket combustion chambers was a small hydrogen peroxide JP-4 rocket engine especially constructed to provide uniform and reproducible combustion. The burner, equipped with a supersonic two dimensional exhaust nozzle, convergent-divergent type, produced a rectangular exhaust gas stream having a cross-section 4 in. × 1 in. After the burner was stabilized, the weighed specimen was inserted into the hot gas stream and exposed for a predetermined time, then the specimen was automatically withdrawn.

The burner can be operated at fuel-oxidizer ratios ranging from stoichiometric, with no excess oxygen, to those producing 20% excess oxygen. Because of the limited chamber pressures available and propellants adaptable for this burner, combustion temperature varied almost inversely with the percentage of excess oxygen. The maximum calculated combustion temperature obtainable with 5% excess oxygen is 4820 R, while at 20% excess oxygen, the maximum is 3260 R. The phenolic asbestos and silicone asbestos specimens were exposed to three gas temperatures having varying percentages of oxygen noted in Table 3.

The asbestos laminate specimens consisted of a ½ in. thick plate 4 in. × 3 in. with a nominal ¼ in. leading edge radius. Fig. 1 illustrates schematically the test apparatus and shows the manner of specimen insertion and removal.

Table 4 summarizes typical results obtained on four of the asbestos laminates exposed to rocket engine exhaust. Analysis of the tests seemed to indicate the higher combustion temperatures of Condition I (4820 R) with 5% oxygen had greater deteriorating effect than the low temperature of Condition III (3260 R) and 20% oxygen.

Based upon the results of Cornell Aeronautical Laboratory tests on asbestos composites, temperature appears to be more critical than the percentage of oxygen in the thermal degradation of these laminates. Due to the amount of material eroded from these specimens, it is doubtful if the asbestos composites evaluated would be satisfactory for fabricating rocket motors or nozzles.

Short-Time Exposure to Ablation Conditions with Varying Heat Fluxes — The equipment used for conducting ablation tests can be considered to be a high-temperature resistance heater capable of introducing up to 65 kw into a nitrogen gas stream. Mass velocities up to 40 lb/ft²-sec are obtainable with choked flow normally existing at the exit. The facility can, at the present time, be used to produce initial calorimetric heat flux densities up to 550 Btu/ft²-sec.

The heart of the test equipment was a specially designed graphite tube containing a fluted core made from graphite forced into its center. The tube and core are heated electrically by a motor generator set capable of delivering up to 1,600,000 watts at 4000 amp d-c. Nitrogen gas flowing in the space between the tube inner wall and the fluted core can be heated to temperatures of approximately 5000 F. Fig. 2 is a sketch of the ablation test facility.

The preliminary experimental work utilized specimens that were injected into the hot-gas stream at angles of 45 and 90 deg to the flow. The specimens were ¼ in. thick, 1 in. wide, and 4 in. long. Four thermocouples were molded in the center of the 4 in. length at intervals of 0.020 in. from the flat surface. These specimens showed that the hot-gas ablated a round crater in the center of the specimen and this appeared to cause the spalled laminate materials to be forced out into the path of the entering gas stream with resultant turbulence.

Other geometric shapes were then evaluated. They consisted of a hemispherical nosed cylinder ½

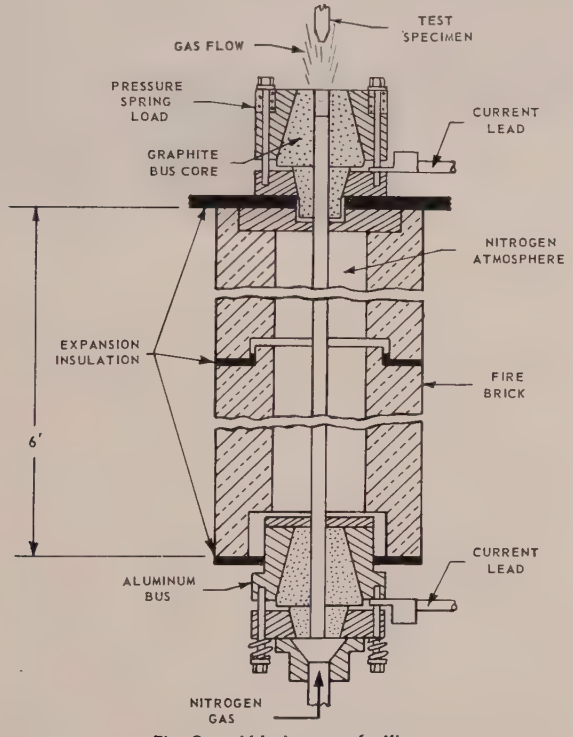


Fig. 2 — Ablation test facility

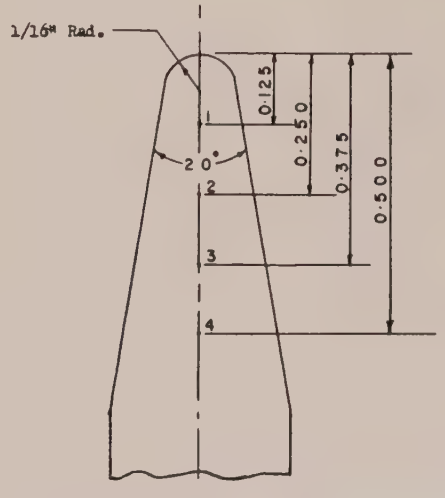


Fig. 3 — Thermocouple placement in wedge specimens

in. in diameter, a ½ in. square specimen with a nose radius of ¼ in., and a 20-deg wedge-shaped specimen ½ in. thick with a 1/16-in. nose radius. From the appearance of the specimens after test, the flow of hot gas appeared to be most uniform on the wedge-shaped specimen. It was then decided to conduct further preliminary ablation studies on this configuration.

Specimens of asbestos laminates were made as a flat plate ½ in. thick, 2¼ in. × 3 in. and machined to shape, since it was felt the thermocouples could be positioned more accurately with this method of fabrication. Four thermocouples of No. 40 iron-constantan wire were molded in each sample. The first thermocouple was positioned on the center line of the center ply, ⅛ in. back from the leading edge. The other three thermocouples were spaced along the same centerline as shown in Fig. 3. All leads were brought out through a 1/16 in. diameter Teflon tube near the trailing edge. The 20 deg wedge-shape with 1/16 in. radius leading edge was machined as a final operation.

Fig. 4 is a view of wedge-shaped specimens after 15-sec exposure to gas temperatures of approximately 4000 F and an initial heat flux of 500 Btu/ft²-sec.

Fig. 5 is a view of the cross-section of the ablated phenolic asbestos wedge-shaped specimen shown in Fig. 4.

A comparison of typical temperature profile of phenolic and silicone asbestos wedge-shaped specimens, as determined by the thermocouples, is shown in Fig. 6.

A comparison of the depth of penetration of the ablation of silicone and phenolic wedge-shaped specimens as a function of time is shown graphically in Fig. 7.

The relative ablation rates of phenolic asbestos and silicone asbestos was found to be similar for initial heat fluxes ranging from 250 to 550 Btu/ft²-sec.

The thermal degradation of the residual nonablated material in the wedge-shaped test specimens

Fig. 4 — Two types of asbestos reinforced plastics after 15-sec ablation test

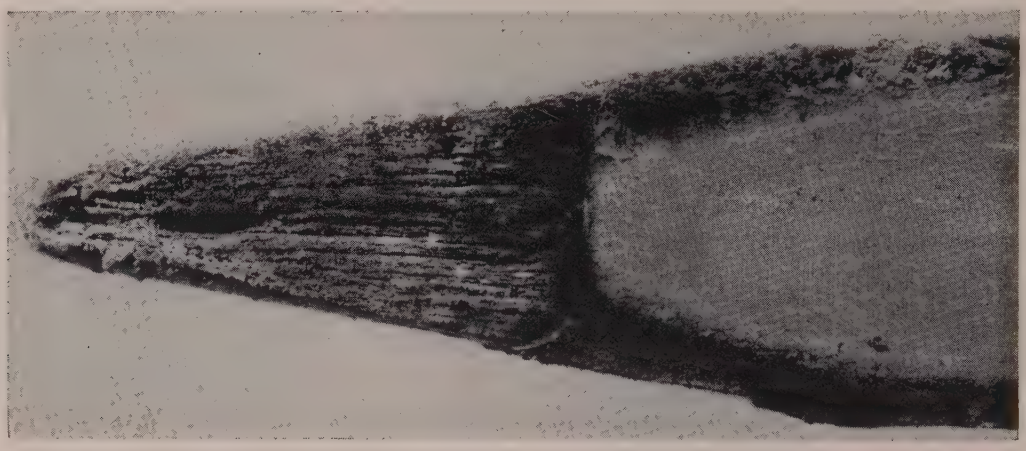


Fig. 5 — Cross-section of ablated phenolic asbestos specimen

Table 5 — Ultimate Compressive Strength Before and After Ablation Tests

Control	Phenolic Asbestos, psi	Silicone Asbestos, psi
Before Test	24,900	12,700
	24,200	12,800
	25,300	13,600
	24,200	14,000
	24,600	13,300
Average	24,600	13,300
After Test	18,400	8,200
	14,100	8,300
	16,400	8,700
Average	16,300	8,400

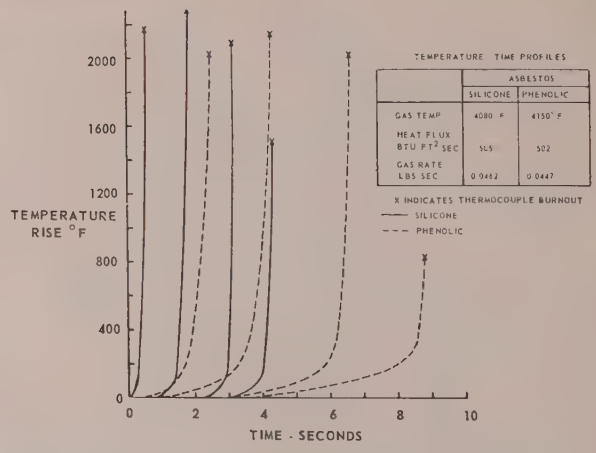


Fig. 6 — Temperature-time profiles for silicone and phenolic asbestos

was studied by sectioning the residual material into small specimens and determining their compressive strength when loaded edge-wise. This procedure required very careful preparation of the specimens and careful handling during the tests.

The edge-wise compressive tests specimens were typically 0.275 in. square and 0.5 in. long. These were cut, using an abrasive wheel, from the center of the ablated region, perpendicular or parallel to the leading edge. Typical comparative results of the loss in compressive strength are given in Table 5.

Long-Time Exposure to Temperatures up to 1000 F — To obtain an indication of the amount of thermal degradation of asbestos laminates, the weight losses after exposure to temperatures ranging from room temperature up to above 1000 F were determined. A comparison of the loss in weight of phenolic asbestos and silicone asbestos laminates after 3-hr exposure are shown in Fig. 8.

The mechanical properties of asbestos laminates

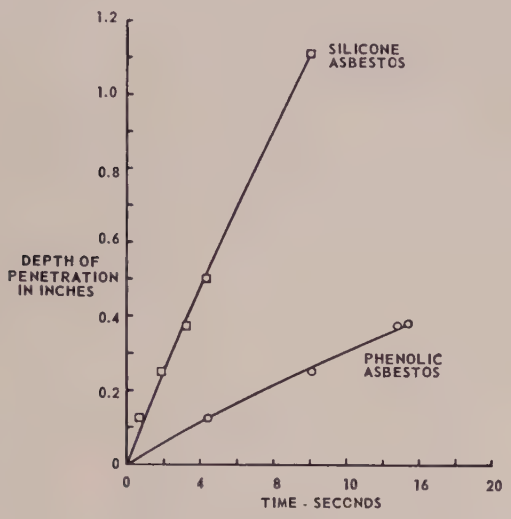


Fig. 7 — Comparison of depth of penetration of ablation front versus time — silicone and phenolic asbestos laminates

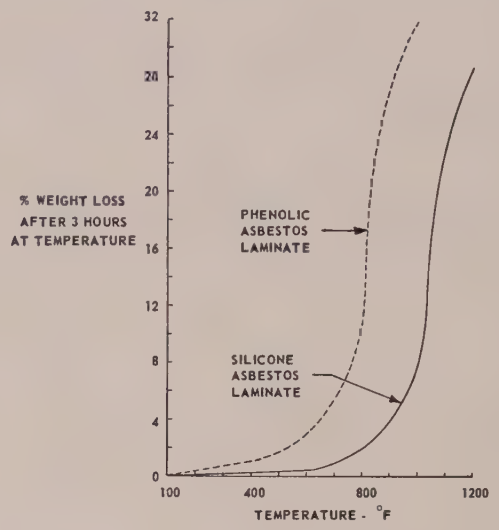


Fig. 8 — Typical weight loss of asbestos laminates, ⅛ in. × ½ in. × 3 in. after 3 hr at temperature

Table 6 — Mechanical Properties of Phenolic Asbestos Laminate Panels at 75 F

Ironsides No. 101 — Asbestos Felt Style 40 RPD

Panel No.	Resin in Laminate, %	Molding Pressure	Thickness Laminate	Ultimate Tensile Strength, psi	Ultimate Compressive Strength, psi	Ultimate Flexural Strength, psi	Flexural Moduli × 10^6
535 A	31.4	15	0.156	14,800	3700	11,800	
B				10,700	3300	11,400	
C				14,500	3300	11,000	1.3
D				14,800	4100	11,300	
E				14,300	4500	10,400	
Average				13,820	3780	11,180	
536 A	32.0	30	0.140	20,300	5600	15,800	
B				20,800	5800	19,900	
C				19,000	5700	17,600	2.0
D				18,900	6400	15,400	
E				19,700	6000	15,700	
Average				19,740	5900	16,880	
537 A	32.0	60	0.115	28,500	11,100	24,700	
B				28,400	10,400	25,700	
C				28,500	10,400	24,500	3.2
D				28,300	10,600	23,200	
E				29,300	11,800	23,100	
Average				28,600	10,860	24,240	
538 A	28.2	200	0.90	57,200	23,200	53,200	
B				56,700	22,000	52,600	
C				60,800	22,700	50,400	4.5
D				57,600	25,100	54,300	
E				51,200	24,800	53,400	
Average				56,700	23,560	52,780	
530 A	27.1	400	0.98	62,100	24,000	55,600	
B				60,500	24,300	57,300	
C				60,800	25,800	58,000	5.4
D				64,600	24,300	56,100	
E				58,700	26,600	53,800	
Average				61,320	25,000	56,160	

molded at varying pressures were determined at room temperature. Results shown in Tables 6 and 7 indicate that for optimum strengths, the molding pressure for both phenolic and silicone asbestos laminates should be in excess of 200 psi.

It was arbitrarily decided that 2% weight loss upon exposure to elevated temperature would be the maximum amount of pyrolysis desirable. Based upon this assumption, it can be observed from Fig. 8 that the corresponding temperature for phenolic asbestos is 500 F and for the silicone asbestos 700–800 F. These temperatures were used for further evaluation of mechanical properties of these two types of laminates. The exposure period was arbitrarily selected as ½ hr and 192 hr. Table 8–10 summarize the results of these tests.

Summary

This paper has reviewed the work carried out in the development of a relatively high-strength, temperature-resistant, composite material and the brief evaluation of its possible uses as material for the fabrication of rocket nozzles, nose cones, rocket bodies, or other rocket components.

This program revealed the fact that when long fibered chrysotile asbestos fibers are used as a reinforcement for heat resistant, phenolic or silicone resins, the composite must be molded at pressures above 200 psi for optimum properties.

Structural laminates molded at pressures of 200 and 400 psi were evaluated in the laboratory to simulate three possible rocket conditions.

The phenolic and silicone asbestos specimens exposed to the first test, simulating rocket motor conditions of high temperature, erosion, and variation in the percentage of oxygen fared rather badly. The percentage of oxygen had no appreciable effect on the rate of degradation. The phenolic asbestos specimens showed considerably lower weight loss due to erosion than the silicone asbestos. However, neither material was considered satisfactory in the present state for use in most rocket motor applica-

Table 7 — Mechanical Properties of Silicone Asbestos Laminates at 75 F

DC-2106 Resin — Asbestos Style 41 RPD

Panel No.	Molding Pressure	Laminate Thickness	Ultimate Tensile Strength, psi	Ultimate Compressive Strength, psi	Ultimate Flexural Strength, psi	Flexural Moduli × 10^6
503 A	30	0.171	12,200	6000	16,700	
B			13,100	6500	18,300	
C			18,900	6100	18,700	1.4
D			18,200	6600	18,900	
E			18,600	6100	17,900	
Average			18,200	6260	18,100	
504 A	200	0.134	23,500	21,100	25,600	
B			24,000	18,600	24,800	
C			24,900	19,200	24,900	2.4
D			24,200	20,800	25,100	
E			24,700	21,600	25,800	
Average			24,260	20,260	25,240	
513 A	60	0.130	20,300	10,800	21,800	
B			20,600	9,600	20,200	
C			19,900	10,200	21,100	1.6
D			20,400	10,700	21,600	
E			21,000	10,100	20,400	
Average			20,440	10,280	21,020	

Table 8 — Phenolic Asbestos Laminates

Ironsides No. 101 — Molding Pressure 400 psi

Mechanical Properties at 75 and 500 F after ½ hr at 500 F

Panel No.	Ultimate Tensile, psi		Ultimate Compression, psi		Ultimate Flexural, psi		Flexural Moduli × 10^6	
	75 F	500 F	75 F	500 F	75 F	500 F	75 F	500 F
530 A	62,100	52,400	24,000	20,200	55,600	39,900		
B	60,500	50,600	24,300	18,400	57,300	41,200		
C	60,800	49,200	25,800	19,600	58,000	42,600	5.4	5.2
D	64,600	48,700	24,300	21,100	56,100	40,500		
E	58,700	51,200	26,600	18,300	53,800	40,300		
Average	61,300	50,420	25,000	19,520	56,160	40,900		

Mechanical Properties at 75 and 500 F after 192 hr at 500 F

Panel No.	Ultimate Tensile, psi		Ultimate Compression, psi		Ultimate Flexural, psi		Flexural Moduli × 10^6	
	75 F	500 F	75 F	500 F	75 F	500 F	75 F	500 F
530 A	62,100	48,600	24,000	13,200	55,600	42,500		
B	60,500	49,200	24,300	14,000	57,300	44,600		
C	60,800	49,800	25,800	12,800	58,000	43,200	5.4	5.0
D	64,600	50,100	24,300	12,600	56,100	45,000		
E	58,700	48,800	26,600	13,700	53,800	42,200		
Average	61,300	49,300	25,000	13,260	56,160	43,500		

tions.

The second test which roughly simulated rocket nose cone conditions, again demonstrated a superiority of the phenolic asbestos over the silicone asbestos composite. Based upon depth of penetration of ablation and the amount of material ablated away, phenolic asbestos was considerably more resistant to ablation. This phenolic asbestos composite merits further review as a material for nose cones.

The third series of tests were concerned with the loss in weight and strength at temperatures of approximately 1000 F for periods up to 192 hr. In this environment silicone asbestos laminates maintain a high percentage of their mechanical properties up to 800 F. This silicone asbestos composite should be considered for further evaluation for rocket bodies or structural components that will be subjected to temperatures of 800 F range for relatively long periods of time.

Table 9 — Silicone Asbestos Laminates
DC-2106 Molding Pressure 200 psi
Mechanical Properties at 75 and 500 F after ½ hr at 500 F

Panel No.	Ultimate Tensile, psi		Ultimate Compression, psi		Ultimate Flexural, psi		Flexural Moduli × 10^6	
	75 F	500 F	75 F	500 F	75 F	500 F	75 F	500 F
519 A	23,200	15,000	11,900	7,620	24,600	17,900		
B	22,900	13,100	13,800	8,900	24,200	17,400		
C	22,500	13,400	12,900	10,600	23,900	18,100	2.5	2.3
D	23,100	15,000	12,700	9,650	24,100	17,500		
E	22,800	15,000	13,000	10,200	23,800	17,800		
Average	22,900	14,300	12,860	9400	24,120	17,800		

Mechanical Properties at 75 and 700 F after ½ hr at 700 F

Panel No.	Ultimate Tensile, psi		Ultimate Compression, psi		Ultimate Flexural, psi		Flexural Moduli × 10^6	
	75 F	700 F	75 F	700 F	75 F	700 F	75 F	700 F
520 A	23,600	12,100	13,800	8740	24,600	16,100		
B	24,100	10,000	12,900	5800	25,000	16,200		
C	23,700	12,000	12,200	7000	23,900	15,900	2.5	2.2
D	24,200	11,300	11,800	7570	24,800	16,200		
E	23,900	11,000	13,100	5240	24,200	16,800		
Average	23,900	11,300	12,760	6870	24,500	16,200		

Mechanical Properties at 75 and 700 F after 192 hr at 700 F

Panel No.	Ultimate Tensile, psi		Ultimate Compression, psi		Ultimate Flexural, psi		Flexural Moduli × 10^6	
	75 F	700 F	75 F	700 F	75 F	700 F	75 F	700 F
521 A	25,200	16,000	15,600	8,430	28,400	19,900		
B	24,900	14,800	16,100	14,200	26,600	21,500		
C	26,000	14,000	14,800	13,700	26,800	19,700	2.9	2.8
D	24,800	16,200	14,100	14,000	27,200	20,900		
E	25,300	15,900	15,300	9,320	27,800	20,600		
Average	25,240	15,400	15,180	11,900	27,360	20,500		

Table 10 — Silicone Asbestos Laminates
DC-2106, Molding Pressures 200 and 400 psi
Mechanical Properties at 75 and 800 F for 192 hr at 800 F[a]

Panel No.	Molding Pressure	Ultimate Tensile, psi		Ultimate Compression, psi		Ultimate Flexural, psi		Flexural Moduli × 10^6	
		75 F	800 F	75 F	800 F	75 F	800 F	75 F	800 F
518 A	200	27,600	9,200	23,200	13,600	28,100	16,200		
B		26,200	10,700	21,100	14,600	29,200	15,800		
C		28,100	12,000	23,200	8,100	28,800	16,000	2.5	2.1
D		25,800	13,300	20,600	8,100	28,600	18,900		
E		26,700	9,000	20,100	7,300	27,900	12,000		
Average		26,880	10,040	21,640	10,340	28,320	15,780		
Strength Retained, %[b]		37		48		58			
529 A	400	29,800	16,000	24,100	16,660	28,900	18,600		
B		30,200	17,100	23,600	14,200	28,500	22,100		
C		31,000	16,500	22,700	17,000	27,800	20,900	2.8	2.6
D		29,100	17,100	21,100	16,700	28,100	17,600		
E		29,900	15,700	22,000	16,800	28,400	17,700		
Average		30,000	16,480	22,300	16,260	28,200	19,380		
Strength Retained, %[b]		55		73		69			

[a] Samples subjected at 800 F for 192 hr in circulating air oven, allowed to cool to room temperature, then held at ½ hr at 800 F in test oven before testing at 800 F.

[b] Percentage of room temperature strength left at 800 F after exposure to 800 F for 192 hr.

RADIATION EFFECTS STUDIES for the development of radiation-resistant aircraft subsystems

J. J. Tierney
Convair Division, General Dynamics Corp.

THE RADIATION ENVIRONMENT of a nuclear-powered aircraft is basically made up of neutron and gamma radiation. Each of these types of radiation has a wide energy spectrum and each has an average energy of about 1 mev. Figs. 1 and 2 show these spectra for the ground test reactor (GTR).

Fig. 1 is a plot of the differential neutron flux versus energy. The plot has been segmented over seven different energy intervals and the integrated differential flux is shown for each interval. Fig. 2 contains plots of both neutron and gamma differential flux. The energy range shown contains those portions of the neutron and gamma radiation which contribute significantly to radiation damage.

Energy Transfer

Neutron and gamma radiation energy transfer to a material or accessory resulting in damage is accomplished in the different processes discussed below.

Neutron Energy Transfer — Neutron interaction with matter is varied and complex. In the study of radiation effects, however, there are only two important types of interactions. These are: (1) elastic scattering and (2) absorption by atoms of the bombarded material, resulting in transmutations of the atoms accompanied by secondary radiations of widely varying types. In general, the damage to a material from transmutation of atoms can be ignored and the secondary radiations are important only in consideration of hazards to personnel.

Elastic collision is a billiard-ball-type collision between a neutron and the atoms of a material, in which energy and momentum are conserved. The average energy transferred to an atom by a neutron in an elastic collision is a function of the incident energy of the neutron and the mass of the struck atom. The greatest transfer of energy is with the lighter elements, that is, those with small mass numbers: the heavier the element, therefore, the smaller the average energy decrement of the neutron per collision.

The energies of the neutrons emitted by a reactor source are sufficient to displace the atoms in a ma-

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 9, 1959.

THIS PAPER has been prepared primarily for the aircraft equipment design engineer who has little or no experience in the field of radiation damage. It represents an effort on the part of the author to acquaint these people with the problems of operating aircraft equipment in the radiation environment of a nuclear-powered aircraft and to indicate the studies being performed in an effort to produce radiation-resistant aircraft subsystems.

Some basic considerations of the radiation environment and its interaction with matter are given, followed by a general discussion of radiation effects on and resultant damage to aircraft materials and accessories. Also discussed are the studies required and being conducted to render solutions to the problems of operating equipment in this radiation environment.*

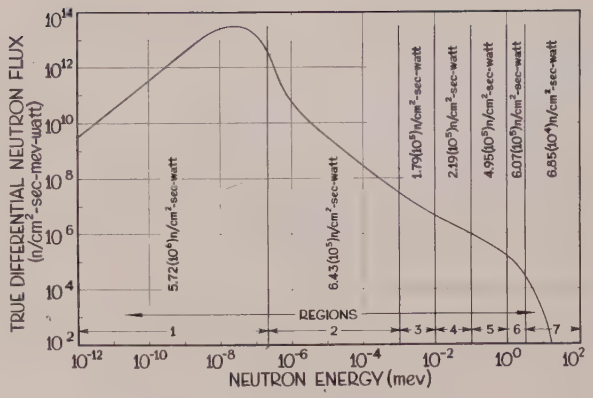


Fig. 1 — GTR neutron spectrum

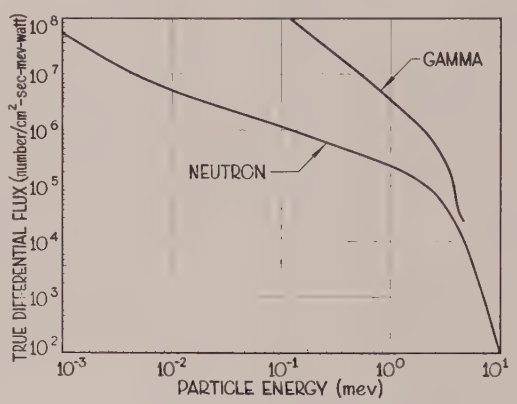


Fig. 2 — GTR neutron and gamma spectra

terial with energies capable of causing secondary displacements and ionization. The extent of damage suffered by a material from neutron bombardment is, consequently, a function of its sensitivity to atomic displacements and/or ionization.

Gamma Energy Transfer — The transfer of energy from gamma rays to a material is accomplished by three different processes: photoelectric effect, Compton scattering, and pair-production. The photoelectric effect is the collision of a gamma ray with an electron, in which there is a complete transfer of energy to the electron, resulting in the disappearance of the gamma ray. The Compton effect is the collision of a gamma ray with an electron, in which there is only partial transfer of energy to the electron; this results in a scattered electron and a gamma ray degraded in energy by an amount equal to the energy transferred to the electron.

In the third process, pair-production, the incident gamma ray is transformed into an electron and a positron (positive electron), a reaction requiring a threshold energy equivalent to two times the rest mass of the electron (M_o) times the velocity of light squared (c^2), or 1.02 mev. The total kinetic energy of the resulting electron and positron is roughly equal to the energy of the incident gamma ray minus the energy required for the transformation, $2M_oc^2$. These particles then undergo collisions with other electrons until they approach thermal equilibrium with the medium, at which time the positron combines with an electron and the two disintegrate into two photons, each having an energy equal to M_oc^2 or 0.51 mev.

The relative importance of these three processes is a function of the gamma-ray energy and the atomic number of the bombarded material. The photoelectric effect is predominant at the lower gamma energies up to approximately 0.7 mev. As the energy of the incident gamma ray increases, the Compton effect becomes predominant — between about 0.05 and 11 mev. As previously stated, pair-production cannot occur below 1.02 mev and becomes increasingly important as the energy of the incident gamma ray increases.

The effect of each of these processes is to ionize the bombarded material. The damage to a material through gamma energy transfer is, then, a function of the sensitivity of the material to ionization.

Chemical Bonds

The relative importance of neutron and gamma radiation in causing damage depends upon the susceptibility of the material to change, either from atomic displacements or from the effects of ionization. The material's susceptibility to damage from either means is determined by the type of chemical bond holding the atoms of the material together. Three types of chemical bonds are considered when classifying materials: metallic, ionic, and covalent.

Metallic Bonds — A metallic bond consists of a lattice of positive ions whose outer-shell electrons are free to migrate. Because of the mobility of the electrons in metals, ionizing radiation has no effect

except to produce heat. On the other hand, the properties of metals are very dependent on the lattice structure; imperfections produced by the displacement of atoms from their normal positions in the lattice can cause significant changes in the properties of the material. When considering the damaging effects of radiation on metals, it is apparent that only neutron radiation is important.

Ionic Bonds — An ionic bond is one in which electropositive and electronegative atoms transfer electrons and the atoms are held together in a lattice of negative and positive ions by electrostatic forces. In ionic-bond materials, ionizing radiation will eject electrons from the ions leaving free atoms or positively charged ions, but not affecting the lattice structure. In most cases, electrons migrate back to the atoms and recombine with them to form the normal state. However, the electrons can be trapped at points of imperfection in the lattice and form what are called "F-centers." These F-centers cause discoloration of the material, but have little effect on their chemical or physical properties. Neutron radiation, on the other hand, causes displacements of the atoms in the lattice structure and consequent changes in the physical properties of ionic-bond materials, as in the case of metals. The relative importance, therefore, of neutron and gamma damage depends on the material and its application. When a material is required to retain its structural properties, the damage produced by neutrons is most important. On the other hand, gamma damage is very important in optical elements. In this instance, the production of F-centers resulting in decreased light transmission makes ionizing radiation most important.

Covalent Bonds — A covalent bond consists of molecules made up of electron-sharing atoms, the molecules being held together by weak van der Waals forces. Organic compounds as well as most common gases and liquids fall into this category. The ionization caused by neutron or gamma radiation is sufficient to break these covalent bonds, producing atoms, different molecular groups, and, sometimes, different chemical bonds. The mechanism of damage in these materials is complex and it is not always possible to predict whether neutron or gamma radiation contributes more to damage. Both forms of radiation must be considered significant until data exist from which the relative importance can be ascertained. Their relative importance is certainly a function of the neutron-to-gamma ratio.

Property Changes After Irradiation — In the study of radiation damage to metallic- and ionic-bond structures, the process of annealing is an important consideration. A distorted lattice is thermodynamically unstable and atoms free to migrate will return to their normal positions in the lattice, removing the effects of neutron-caused displacements. The degree of annealing depends upon the material and its temperature. Some annealing occurs in most metals at room temperature, whereas in materials such as ceramics, elevated temperatures are required before significant annealing occurs. Also of importance is the possibility of chemical changes continuing to take place in covalent-bond materials after irradiation. This could occur if, in the material under irradiation, the rate of certain chemical reactions occurring as a result of irradiation were a function of the number of free radicals being formed during irradiation. Such reaction rates can be strongly dependent upon temperature.

It is apparent that, if these conditions exist, property changes recorded after irradiation could possibly be a function not only of the irradiation conditions, but of the storage temperature of materials after irradiation and the time after irradiation at which the measurement is made.

Radiation Effects

The effects of radiation on an aircraft material or accessory may be transient or permanent in nature. Transient effects are those which exist in the presence of a radiation environment, but disappear in its absence. Permanent effects are those changes which occur under irradiation and persist after irradiation.

Transient Effects

Transient effects are rate-sensitive phenomena which become more significant with increased radiation rates. The magnitude and/or significance of transient effects are time independent, except in cases where they are altered or initiated by permanent radiation effects. Some of the transient phenomena which have been observed and which might constitute serious problems in a nuclear-powered aircraft are: (1) ionization of air and formation of ozone, (2) radiation heating, (3) gas evolution in organic materials, and (4) transient effects in electronic parts.

Ionization of Air and Formation of Ozone — The ionization of air, which occurs even at moderate flux levels, should, perhaps, not be considered in all instances as a problem in itself, but more accurately as an unfavorable radiation-induced environment which seriously affects the functioning of certain electrical and electronic components.

Our experiments have afforded ample evidence that air ionization significantly reduces the effectiveness of insulation resistance. However, in the experiments thus far, it has not ben possible to determine exactly the magnitude of ionization. At high flux levels, the probable effects of air ionization on electronic components can, theoretically, be extended to include problems in high-voltage breakdown, excessive relay arcing, erroneous thyratron, triggering, voltage-level errors in gas-filled VR tubes, and the like. The latter two, of course, pertain to ionization of gases other than air.

Ozone is produced in an all oxygen-containing atmospheres under irradiation. The concentration is strongly dependent upon the radiation rate and temperature of the atmosphere. Ozone is a problem insofar as it accelerates the corrosion of metals and reacts chemically with all exposed surfaces of organic materials. The amount of ozone damage sustained by organic materials is also a function of the stress condition of the material. Damage to organic materials irradiated in air has been observed to be more severe than for the materials irradiated in an inert atmosphere.

Ozone formation can also present a hazard to personnel if formed in crew environmental systems. The maximum allowable concentration is set at 0.04 parts per million and concentrations as high as 0.159 ppm have been observed in the irradiation of a liquid-oxygen supply system at a radiation rate

of 7×10^3 ergs/gm(C)–hr for gamma and 4×10^7 n/cm^2–sec ($E > 0.4$ ev) for neutrons. However, because of its chemically active nature, removal of ozone from crew-breathing supply systems does not present too much of a problem.

Radiation Heating — At high radiation rates, the heat generated in materials by neutron and gamma energy absorption can cause considerable damage to system parts and components unless proper provisions are made for cooling. The heat generated is a function of the radiation rate incident on the material, the density and geometry of the material, and the absorption cross-sections of the material. Heat dissipation is not a new problem in design, but the requirements for the dissipation of radiation heating are more extensive than those previously encountered.

In the presence of a high-radiation field, all materials become heat sources. This means that structural materials, fuel, wiring, plumbing, and the like, heretofore of little concern, must now be considered. The problem differs from aerodynamic or solar heating where the heat is generated at the surface of an aircraft and can, in many areas, be isolated from other portions of the structure by insulation. In nuclear-radiation heating, the entire aircraft is subject to heating in direct proportion to the intensity of the radiation.

The heating can be controlled only by controlling the intensity of the radiation, that is, by shielding the source of the radiation. To shield out all radiation completely, however, would impose a prohibitive weight penalty, so that some compromise must be made. In aircraft design, the entire aircraft must be considered and, in evaluating heat-load requirements, necessary provisions must be made for heat removal.

Gas Evolution in Organic Materials — The evolution of gas, which occurs in all organic materials, is not a pure transient phenomenon. It does not necessarily occur immediately upon application of a radiation environment, but apparently begins after a material has been subjected to a certain radiation dose which varies with the material. Because significant gassing occurs only during irradiation and is dependent upon the radiation rate, it is included here for discussion.

Solid organic materials have many applications in present-day aircraft; they are used as structural members, seals, diaphragms, and insulation coatings, to name a few. Most solid organic materials located in an area where the dose reaches 10^{10} ergs/gm(C) will evolve gas. After a material has started to gas, the rate of evolution increases with the radiation dose rate. Many elastomers evolve hydrogen gas when irradiated from 10^{10} to 10^{11} ergs/gm(C); however, in some cases, halogen-containing polymers, such as Teflon and Viton A, may evolve corrosive gases such as chlorine and fluorine.

Explosive gases could build up rapidly in a relatively unventilated area, thereby becoming hazardous. A curve showing the amount of gas evolved from certain solid organic materials as a function of the energy absorbed per gram of the particular material is presented as Fig. 3.

Organic fluids such as hydraulic fluids, lubricating oils, gyro flotation fluids, damping fluids, and those used in electrolytic application present the most adverse problems when irradiated. Gases evolved include CO, CO_2, H_2, and hydrocarbons. Most hydraulic fluids exposed to doses as low as 10^9 ergs/gm(C) evolve enough gas to cause foaming. Lubricating oils and oils used in electrolytic applications evolve gas at doses lower than 10^9 ergs/gm(C). Oils used in sealed oil-filled capacitors have evolved sufficient amounts of gases when irradiated to rupture the sealed cases. The evolution of gases in gyro flotation and damping fluids has been noted at gamma doses of 10^9 ergs/gm(C), contributing to gyro malfunctions.

Transient Effects in Electronic Parts — Transient effects in the operational characteristics of electronic components have been investigated by Convair-Fort Worth using the prompt-burst critical assembly, Godiva II, of Los Alamos National Laboratory. Fig. 4 shows the average fission rate during a Godiva II prompt burst.

The dose accumulated by each component during the burst was approximately 10^{11} fast neutrons ($E > 2.9$ mev), or 10^{12} neutrons ($E > 0.33$ mev). The gamma dose was 4.4×10^5 ergs/gm(C). The total irradiation time was 300 microsec.

Experiments utilizing the Godiva II facility have yielded data indicating significant transient changes in transistors, photo diodes, insulation resistance and/or air ionization, and capacitance. Similar experiments conducted by International Business Machines Corporation (IBM), Owego, New York, showed similar transients in transistors and capacitors, as well as in ferroelectric crystals and diodes.

The results of these investigations, along with supporting evidence from Convair's Ground Test Reactor (GTR) experiments, are discussed below. Included are brief interpretations of the practical effects of such transients on subsystems in which these parts are used.

Capacitors—Both Convair and IBM obtained photographs of capacitance response during a Godiva II burst. Significant changes occurred in the measurements made by both; however, the changes could

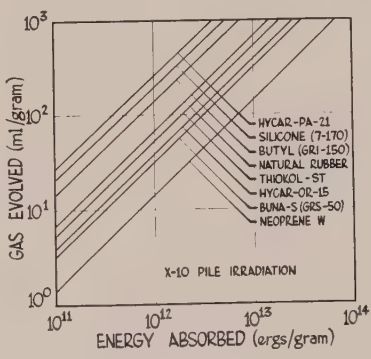


Fig. 3 — Gas evolution for elastomeric materials

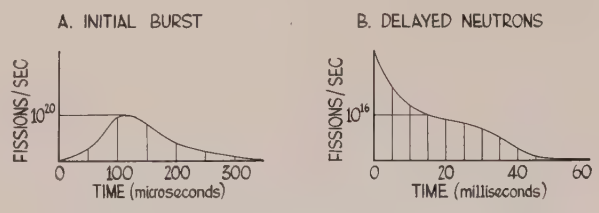


Fig. 4 — Average fission rate during Godiva II prompt burst

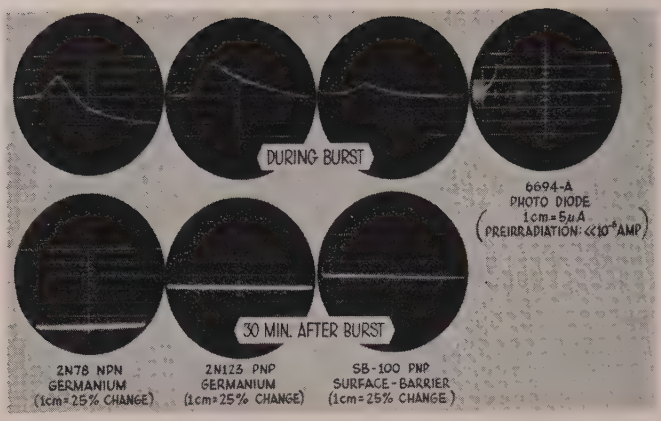


Fig. 5 — Transient responses

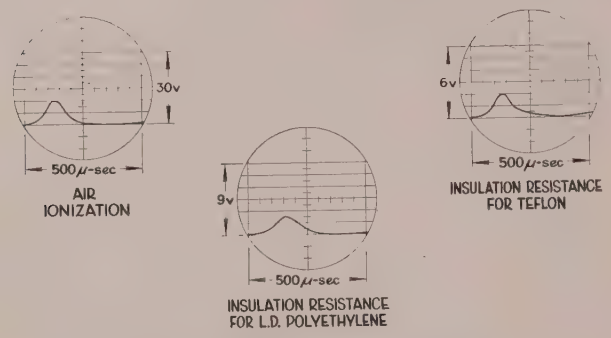


Fig. 6 — Current conduction response during Godiva burst (drop across 7500 ohms with 600 vdc supply)

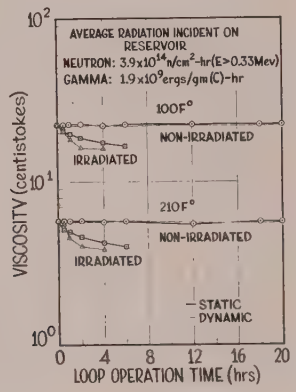


Fig. 7 — MIL-L-7808C viscosity

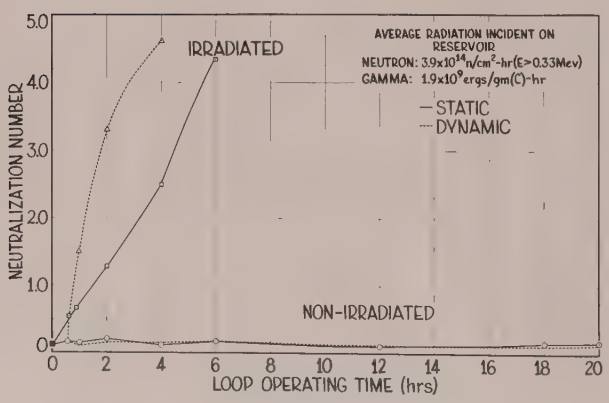


Fig. 8 — MIL-L-7808C neutralization number

not be conclusively attributed to either changes in capacitance or dissipation factor. The general opinion is that insulation breakdown, either dielectric or air, was primarily responsible for the changes in the capacitance measurements. Data from GTR experiments also have shown transient effects in capacitors, the most significant effect being the inability of storage capacitors to retain an electrical change at even moderate dose-rate levels. High power dissipation might also be a problem.

Transistors and Diodes — The transistor responses pictured in Fig. 5 represent changes in the d-c output current level with a constant d-c input current during a Godiva II burst. The dose-rate sensitivity is evidenced by the close correlation between the responses and the burst pulse pictured in Fig. 4. The response shown reflects changes in leakage current (I_{ceo}) and/or d-c current gain (beta). IBM data which include similar measurements on the leakage current only indicate that most of the transient increases shown in Fig. 5 are due to the leakage component of the d-c output current level; their data showed leakage current increased to 12 milliamp during the burst peak. Transient excursions from a few microamp to 3 milliamp were common according to IBM findings.

As would be expected from transistor I_{ceo} data, silicone and germanium diodes experienced similar transient breakdown in reverse characteristics. A slight transient change in zero breakdown voltage was observed also by IBM.

Transistor trigger circuits have been observed to become unstable in high flux fields. These circuits were conventional, and in all probability equivalent circuits could be designed with some degree of improved radiation stability.

The problems created by diode transients vary from insignificant to serious, depending on the applications. Diodes used in computers would probably be the greatest problem area.

Photosensitive Devices — Fig. 5 includes a dark-current response curve for a cadmium sulfide photo diode. The magnitude of the response at the peak of the burst cannot be determined, but can be judged from the photograph to be significant. The recovery time is also unknown, but the diode did recover to approximately the pre-irradiation value.

Results of a GTR experiment show transient effects on RCA 5819, DuMont 6292, and RCA 1P21 photomultiplier tubes. The dose-rate sensitivity varied among tube types. At maximum fluxes of 3×10^9 N_f/cm²-sec (> 2.0 mev) and 6×10^6 ergs/gm(C)-hr (gamma), dark currents increased from fractions of microamperes to 12, 35, and 50 microamp for the 5819, 1P21, and 6292 type photomultiplier, respectively.

The results illustrated indicate transient damage which would be intolerable in many critical applications. Even moderate flux levels reduce the signal-to-noise ratio significantly, and indications are that high flux levels will render many photosensitive devices useless.

Insulation Resistance — Separation of an insulating material's normal conductivity from the conductivity due to air ionization is not readily accomplished. Fig. 6 shows photographs of current conduction through air, polyethylene, and Teflon during the Godiva II radiation burst. Certainly, a significant part of the current through the poly-

ethylene and Teflon was due to air ionization; however, the difference in current value between polyethylene and Teflon, both having the same geometry, is evidence of some conduction via the insulating materials.

This problem presents obvious difficulties in low-level instrumentation and high-voltage equipment. Some improvement may be found in pressurization or evacuation of such equipment.

Permanent Effects

Permanent effects caused by irradiation which have been observed and which might constitute serious problems in a nuclear-powered aircraft are radiation damage to: organic fluids, organic solids, electronic parts, metals, and other materials.

Damage to Organic Fluids — Depending upon the fluid and the environment, permanent damage to organic fluids most often exhibits itself in one or more of the following forms: a change in the viscosity, a radical increase in the neutralization number, a decrease in oxidation stability, a decrease in lubricity, flash point decreases, particle formation, and gelation. Although some of these changes occur in the absence of a nuclear environment, nuclear radiation serves to accelerate these changes, making the damage considerably more severe for a given period of service.

Lubricating Fluids — An irradiation test on a MIL-L-7808C lubricating fluid was conducted in our laboratories in which a simple test loop was constructed for the purpose of obtaining dynamic data. The loop consisted of a motor-driven pump, a pressure control valve, a filter, a relief valve, and a reservoir with provisions for controlling temperature and drawing off fluid samples during irradiation. The loop was operated at 60–70 psi and bulk-oil temperatures were maintained at 300 F. To obtain information on the difference in data obtained in static irradiations from that obtained in dynamic irradiations, a second reservoir was placed in the panel and filled with fluid to be irradiated statically. This fluid was also maintained at 300 F. To establish fluid property changes of the non-nuclear environment, a pre-irradiation run was made under the same conditions that existed during irradiation. By comparing these data to irradiation data, the contribution made by the radiation environment was determined.

Fig. 7 is a plot of viscosity versus pre-irradiation and irradiation-run time for both the static and dynamic samples. It can be seen that the viscosity decreased more rapidly during irradiation and that changes in the dynamic samples exceeded those in the static samples. In Fig. 8 the same information is presented for increases in neutralization number. Radical changes in the neutralization number occurred, and again it is apparent that the dynamic sample degraded more severely than the static sample. Other property changes sustained by the fluid include decreased oxidation stability, loss of lubricity, and increased precipitation number (indicating particle formation). In Figs. 7 and 8, no data are shown beyond the 5-hr point for the irradiated samples. This is because the fluid had degraded far beyond tolerable limits at this time and the planned 20-hr irradiation was terminated. At the time of termination, the radiation which the fluid had received was approximately 9×10^9 ergs/gm(C) of gamma and 2×10^{15} neutrons/cm^2 ($E > 0.33$ mev). Such property changes as these in a fluid in use in an engine oil system or in some component lubricating system would rapidly manifest themselves in the form of excessive bearing and gear wear, internal corrosion and filter clogging, resulting in failure of the lubricated accessories.

Hydraulic Fluids — Operating hydraulic-fluid loop irradiations, employing such fluids as MLO–8200, OS–45, MIL–L–5606, and Oronite 8515, have been conducted in subsystem tests. The radiation levels at which these fluids began to degrade varied from 10^9 ergs/gm(C) for gamma and 5×10^{14} n/cm^2 ($E > 0.33$ mev) to 10^{10} ergs/gm(C) and 5×10^{15} n/cm^2 ($E > 0.33$ mev).[1] Degradation in general followed the same pattern as in lubricating fluids. Among the effects noted were changes in viscosity, decreased oxidation stability, increased neutralization numbers, decreased lubricity, and particle formation. The resultant effects of these property changes in an aircraft hydraulic system include excessive wear of moving parts, internal corrosion, and filter clogging. Also, particles formed could cause the plugging of small orifices in such control elements as an electrohydraulic servo valve resulting in flight-control malfunction. The fluid property changes in combination with gases evolved could also cause the response characteristics of a hydraulicaly-powered flight control system to alter sufficiently to degrade system performance or cause the system to fail.

Others — Other applications in which organic fluids are used in an aircraft include gyro flotation and damping fluids, accelerometer damping fluids, and damping fluids or sensing media in other types of transducers. In the irradiation of gyro assemblies at Convair, fluorolube flotation fluid degradation has resulted in gyro failure at radiation levels of about 5×10^9 ergs/gm(C) gamma and 2×10^{15} n/cm^2 ($E > 0.33$ mev). Fluid damage resulted in decreased viscosity, increased acid number, particle formation, and gelation in narrow fluid passages. Failure of the components was marked by excessive drift and gyro bottoming.

Fluids in common use as damping fluids have been observed to gel above 5×10^9 ergs/gm(C) gamma and 2×10^{15} n/cm^2 ($E > 0.33$ mev). Increasing viscosities will cause significant changes in damping characteristics and signal errors before these radiation levels are reached. The function of these components which utilize fluids is that of providing accurate information on a given parameter. This information is either transmitted to the pilot for purposes of controlling the aircraft or monitoring the system operation, or it serves as a signal necessary to the proper functioning of any one of many automatically controlled operations throughout the aircraft. The obvious effects of errors in these functions are intolerable.

Damage to Organic Solids — Permanent damage to organic solids appears in the form of many different property changes. The relative importance of these changes is, as always, a function of the application of the material. The more notable changes are: an increase followed by a decrease in tensile strength, a decrease in elongation, embrittle-

[1] Values such as these in this paper represent approximate values. Exact values are dependent upon the property of interest and the irradiation conditions.

ment, an increase in compression set, a decrease in peel strength, and a decrease in dielectric strength.

Seals — Solid organic materials in many different forms serve as seals in fuel, oil, hydraulic, and air-conditioning systems. These forms include O-rings, back-up rings, packings, valve seats, tank sealants, and diaphragms. The latter also serve in pressure sensing and control functions. In subsystem tests of these types at Convair, radiation damage to these materials has been apparent in seal extrusion, permanent set, embrittlement, chipping, and flaking. This damage has resulted in internal and external leakage leading to control-valve and relief-valve malfunctions, loss of pressure, loss of fluid, and ultimate assembly failures. Radiation-induced damage to seal materials in common use becomes significant over a wide range of radiation levels, from as low as 5×10^8 ergs/gm(C) gamma and 3×10^{14} n/cm^2 ($E > 0.33$ mev) to as high as 10^{11} ergs/gm(C) gamma and 5×10^{16} n/cm^2 ($E > 0.33$ mev). Plastics in many cases exhibit greater radiation resistance than elastomers, notable exceptions being Teflon and nylon.

In obtaining radiation damage information on organic materials for a given application, it is important that the medium in which they are to be used be considered. Table 1 illustrates the effect of the medium on an O-ring compound, Stillman EX511C, irradiated in air and in JP–4 fuel. Each value shown represents an average value for 5 samples. Control data on the air samples are "as received" values and those for the fuel represent values obtained after a 14-day soak in JP–4. The conditions for the samples irradiated in fuel were the same as for the control samples, except for irradiation. It can be seen that for the sample irradiated in air, tensile strength increased and per cent elongation severely decreased. In the case of the samples irradiated in JP–4, tensile strength decreased and per cent elongation again severely decreased. The dependence of damage upon the medium in which a material is irradiated is an important consideration in the study of radiation effects on all materials.

Electrical Insulating Material — Another important application of organic solids in aircraft is as an electrical insulating material in electronic equipment. Permanent damage to insulating materials is characterized by decreased resistivity, embrittlement, cracking, and flaking. The radiation level at which this damage becomes important is again a function of the particular material and its application and falls in the same range as that given for seal materials. Severe physical property degradation precedes significant electrical property changes. The results of insulation deterioration of this type are obvious and will not be elaborated upon except to say that electronic equipment failures have been observed which are solely attributable to insulation failure.

Others — There are numerous other applications of organic materials in aircraft, including tires, protective coatings, adhesives, shock mounts, and vibration isolators. Statements concerning changes in properties in these would only be a reiteration of what has already been said. The results of deterioration in these materials can readily be realized from the nature of their functions.

Damage to Electronic Parts — Electronic parts fall into many functional categories, each containing many different types. Parts in the more basic functional categories are presented in table form. Table 2 indicates the category, the characteristics of failures, the range of doses at which these failures occur, and lists some types in order of increasing resistance to radiation. Damage sustained by a given type of part has been observed to vary appreciably with the manufacturer. This indicates a definite need for a study of the materials and processes used in the manufacture of parts to determine which yield more radiation-resistant parts. More than ever, the requirement is felt for strict controls over materials and processes as well as for well-established quality control programs to insure uniformity of products. The manner in which changes in an electronic part affect a given piece of equipment is determined by the function and criticality of that part in the circuit in which it is used. Discussion of the effects on the numerous and complex circuits in use is beyond the scope of this paper; however, Fig. 9 is a typical example of component degradation as a result of degradation of its constituent parts. The degradation of the amplifier output voltage was due to the deterioration of two silicon-junction transistors (TI-951) used in the unit. When these were replaced after irradiation, the amplifier performed satisfactorily.

Damage to Metals — Radiation damage in metals is characterized by increased tensile strength, decreased per cent elongation, increased hardness, and decreased ductility. Significant changes in these properties are not observed below an integrated fast-neutron flux of 10^{18} neutrons/cm^2. The degree of change is dependent upon the type of metal, the irradiation temperature, and the initial condition of the metal. At still higher integrated fluxes, $> 10^{19}$ n/cm^2, property changes occur in creep rate, corrosion rate, dimensions, and thermal conductivity..

Damage to metals requires such high radiation levels that it is of concern only in the immediate vicinity of the reactor. It does not require considera-

Table 1 — Influence of Irradiation Media on Stillman EX511C

Medium	Tensile Strength		Elongation, %		Radiation Level	
	Control	Irradiation	Control	Irradiation	Gamma ergs/gm (C)	Neutron (> 2.9 mev), n/cm^2-sec
Air	1817	2057	294	57	1×10^{10}	2.5×10^{15}
JP-4	914	549	209	38	1×10^{10}	2.5×10^{15}

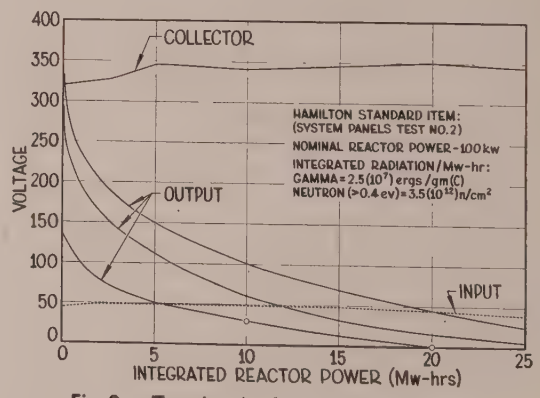


Fig. 9 — Transistorized a-c amplifier damage

tion in other aircraft areas except for activitation of the metal by neutron absorption. Activated metals in the aircraft represent the most significant contributors to ground-handling problems. Where a variety of metals are available for a given application, the selection should be based upon consideration of the neutron absorption cross sections of the material.

Damage to Other Materials — Other materials to be considered are glass, ceramics, and salt compounds. Glass can be found in applications such as filler for organic materials, lenses for lights and light-sensing devices such as star trackers, and observation windows. As a filler for organic materials, damage to glass can be ignored as compared to damage to the organic material. In the other applications mentioned, the darkening of glass caused by F-center formation which results in decreased light transmission is of greatest concern. Significant degradation in TV-picture quality due to lens darkening was noted at 3.1×10^9 ergs/gm(C) gamma and 3×10^{14} n/cm^2 ($E > 0.48$ mev).

Ceramics find wide application as electrical insulating materials in aircraft. As can be seen in the previous section on electronic parts (Table 2), ceramics are far superior to organic materials in radiation resistance and will most likely be employed wherever their properties will allow their use in a given application.

One application of salt materials can be found in fire-detection systems. Knowledge of the melting temperature of a salt and of the fact that the conductivity of the salt increases when passing into the liquid state is used to provide a signal indicating an overtemperature condition in various aircraft areas. Such devices have functioned satisfactorily up to

Table 2 — Electronic Parts Damage
(order of increasing radiation resistance)

Category	Integrated Fast-Neutron Flux	Type	Characteristic Changes
Light Detectors	10^{12}	Photomultipliers	Rate sensitivity predominates
		Lead sulfide cells	Decreased sensitivity followed by annealing to normal in 6 days
Semiconductors	10^{12}–10^{14}	Silicon transistors	Decreased current gain, increased leakage current
		Germanium transistors	
		Germanium diodes	Increased forward resistance, decreased reverse resistance
		Silicon Diodes	
Insulators	10^{12}–$>10^{16}$	Teflon	
		Nylon	
		Silicone rubber	
		Polyvinyl chloride	
		Silicone Resins	
		Formex	Physical property degradation precedes electrical property degradation
		Polystyrene	
		Polyethylene	
		Mylar	
		Mycalex	
		Supramica	
		Phenolic (bakelite)	
		Ceramic	
Resistors	10^{13}–$>10^{16}$	Fixed composition	Primarily, decreased resistance with some increases noted
		Deposited carbon	
		Boron carbon	
		Carbon film	
		Wire wound	
Relays (time delay only)	$>10^{14}$	Pneumatic type	Increased time delay
		Fluid-damped type	
Capacitors	10^{14}–$>10^{16}$	Tantalum electrolytic	Decreased capacitance, decreased leakage resistance, decreased, and increased dissipation factor
		Aluminum electrolytic	
		Paper-solid impregnant	
		Paper-oil impregnant	
		Glassmike	
		Molded mica	
		Ceramic	
Piezocrystals	$>10^{16}$	Barium titanate	
		Quartz	
Vacuum Tubes	$>10^{16}$	Receiving type	No damage observed in Convair irradiations to 10^{16}. However, Admiral reports glass envelope cracking. Differences possibly due to higher thermal fluxes or temperatures.
Gas-Filled Tubes	$>10^{16}$	Voltage regulators	
Electromagnetic Devices	$>10^{16}$	Transformers, chokes, motor-type devices, relays (not time delay)	

NOTE: All doses are for neutrons with energies > 2.9 mev (n_f/cm^2)

Associated gamma doses:

Gas-filled tubes $\approx 9 \times 10^{-5} \frac{\text{ergs/gr (C)}}{n/cm^2}$

All others $\approx 8 \times 10^{-6} \frac{\text{ergs/gr (C)}}{n/cm^2}$

4×10^{10} ergs/gm(C) gamma and 3×10^{14} n/cm^2 ($E > 6.1$ mev.)

Radiation Effects Studies

After this very general look at the effects of radiation and the many ways in which it might affect the performance of an aircraft, the problems may seem insurmountable. However, this conclusion cannot be supported with the information presently available. In the previous discussion, no consideration was given to the aircraft configuration. The difficulty of the problem for any given configuration can be ascertained only by answering the following questions:

1. What are the locations of the components or systems in the aircraft?
2. In this location, what are the nuclear and nonnuclear environments?
3. What are the materials or parts available for use in these components and systems?
4. With these facts established, what are the radiation effects which might be expected in this environment?

To provide all the answers to this last question and to obtain solutions to the problems indicated, radiation-effects studies of the following types are required: fundamental studies, material and electronic-parts damage studies, and subsystem damage studies.

Fundamental Studies

Both theoretical and experimental studies are needed to understand the mechanisms of radiation damage. It is only through this understanding that any systematic approach can be adopted for the development of the radiation-resistant materials and parts required in applications where existing materials and parts prove to be inadequate.

Material and Electronic-Parts Damage Studies

Materials and electronic-parts damage studies serve three purposes: screening existing materials, providing design data on the more promising materials, and supplying information necessary to the understanding of environmental interactions.

A tremendous number of materials and parts exist and are continuously becoming available for use in aircraft. Ideally, radiation resistance of all these should be investigated; however, a detailed study of radiation effects on all these materials and parts is not only impractical but impossible. Therefore, screening tests are necessary to obtain a practical number with which to work. These screening tests are conducted at ambient conditions and to radiation levels which result in detectable damage; little consideration is given to other environments. By analysis of the results, a preferred list of materials and parts which might be of use in a given application can be made.

This preferred group is then considered worthy of further investigation, and design-data tests are initiated. The value of the design data is in proportion to the number of points filled in on the curves that describe the property or characteristic changes as a function of the nuclear and nonnuclear parameters and their combinations. The dependence of damage on radiation dose, radiation rate, radiation energy, neutron-to-gamma ratio, and interactions of nuclear and nonnuclear environments is not always known. Consequently, to obtain each curve representing property measurements during irradiation and under a specific set of environmental conditions, a separate test must be run. Static-material tests, on the other hand, require a separate run to obtain each point on the curve since each point represents a property change for a given dose. It is readily apparent that the completeness of the data is limited by the number of materials and parts investigated and the manpower and facilities available. The best approach, then, is to first generate data representing damage at the environmental limits anticipated in a given material or part application. Further reduction in the list of materials and parts under consideration will most likely result. Additional data can then be obtained on those remaining, as needed.

To provide information necessary for a better understanding of design-data curves and for accurate interpolation of data between the curves, an environmental studies program is required. Environmental studies are equally vital to the preparation of nuclear test specifications. In evaluation or qualification testing of aircraft components and systems, environmental conditions during irradiation must be specified which will result in conservative assessments of radiation damage that might be sustained in the intended aircraft application. To accomplish these goals an environmental studies program must include consideration of the following nuclear parameters:

Neutron dose	Gamma dose
Neutron dose rate	Gamma dose rate
Neutron energy	Gamma energy

Neutron-to-gamma ratio

The studies must also include the nonnuclear environments which are suspected of interaction with these nuclear parameters. Those most often considered are:

Temperature
Atmospheric pressure
Atmospheric content or surrounding media
Other environments giving rise to stress and strain

All of these environments will not make significant contributions to damage in all cases. However, none should be deleted from consideration unless the deletion is based upon sound theoretical or experimental evidence. Most of the ways in which these studies can proceed involve large sample sizes and many replications of a given experiment if they are to yield statistically significant results. It is not possible to include all candidate materials and parts in these studies; therefore, it is necessary to select representative materials which are chemically similar to the "most probable for use" candidate materials. As for electronic parts, the most promising of each type must be selected.

Subsystem Damage Studies

The study of the combined effects of environments is very complex even without the addition of a nuclear environment and has never resulted in simple solutions. Consequently, the need for mocking up the anticipated environment of the aircraft has long been desirable to give greater confidence in determining the suitability of equipment for use on an aircraft. This is accomplished in subsystem damage studies.

Subsystem assemblies are designed and con-

structed to simulate, insofar as is possible, their operation in the aircraft; they are instrumented and operated before, during, and after irradiation. Data accumulted before irradiation are used to establish the criteria for normal operation. By comparing data taken during and after irradiation with the pre-irradiation data, deviations from normal operation are detected. After the completion of the test, the components are dissasembled and the materials and parts examined for evidence of radiation damage. It is at this level that efforts at our company are now being made to develop radiation-resistant subsytems, making use of information presently available from the studies previously discussed.

Convair Subsystem Development Program

In an analysis to determine the susceptibility to radiation damage of an aircraft subsystem, it becomes apparent that there are three courses of action available for improving radiation resistance of an existing or proposed design. These are material substitution, redesign, and relocation or partial shielding.

In a given application, it is often possible to improve the radiation resistance of a component simply by replacing existing materials with more radiation-resistant materials. In other cases, because of a lack of suitable substitute material in a particular design or because of the presence of some radiation-induced phenomenon, component redesign is necessary. As a last resort, it is sometimes possible, in the absence of suitable replacement material or design "fix," to relocate or locally shield a critical item.

Logically, the first thing which should be attempted is material substitution, the most direct and least costly procedure. This is the approach that is presently being pursued at Convair in an effort to produce radiation-resistant subsystems. Basically this approach proceeds as follows. As a result of a configuration study critical components are selected for testing. An analysis of these components is then made to determine the requirements for modification. The components are procured, modified, and built into operating assemblies for testing. An analysis of the results is used to establish requirements for further material substitution or redesign.

Component Selection

The subsystems now under consideration were selected from the five major aircraft system areas in which our company has prime responsibility for the development of radiation-resistant subsystems. These are airframe, secondary power, environmental control, flight control, and flight and engine instrumentation. The seven test assemblies presently being studied are:

1. Hydraulic subsystem.
2. Fuel subsystem.
3. Oil subsystem.
4. Fuel quantity-measuring subsystem.
5. Gyro and accelerometer panel.
6. Air-conditioning subsystem.
7. Tire and brake assemblies.

These do not represent all the system areas of interest to us and future tests will be extended to include other areas.

Analysis

The analysis of the subsystem components is based upon a study of the materials used in the construction of the components. Unfortunately, it is not always possible to obtain satisfactory material breakdowns of components—either because the vendor is not always willing to release this information or because he does not have the information readily available or, in some cases, does not know exactly what materials have been used.

When material breakdowns are available, the requirement for modification is determined by evaluating the radiation-damage data available on the materials used and on the possible alternate materials. Possible alternate materials are at first selected primarily because of their demonstrated greater resistance to radiation. It is then necessary to determine if any of the alternate materials also fulfill the necessary nonnuclear environmental requirements for the application. Based on these considerations, recommendations for modifications are made.

When material breakdowns are not available, it is necessary to disassemble the component and attempt to identify materials visually. If the material can be identified, the above procedure is followed. If the material cannot be identified, a suitable material is selected and the unknown material is replaced with one of known radiation resistance.

Existing data on promising alternate materials is sometimes insufficient. If it is possible, data on these materials are obtained in an irradiation period prior to the testing of the subsystem in which it might be used.

Modifications

Modification of the selected components has sometimes presented problems. Although most modifications of components can be accomplished at our plant, some can be accomplished only at the manufacturers facility. This is because we do not have all the special tools and equipment needed for component tear down and overhaul. In several such cases, vendors have cooperated generously and performed the desired modifications. In other cases, components, by nature of their design, cannot be disassembled for modification and it is necessary that the modifications be effected during fabrication.

Testing

Assembly and Pre-Irradiation — After modification and prior to buildup of the assembly, critical components are examined and subjected to tests that will describe their before-irradiation characteristics. These data are compared with data obtained in identical tests after irradiation.

The components are then mounted in a suitable frame in a manner simulating the operation of the assembly in the aircraft. In positioning the components in the panel, consideration is given to the variable fluxes the individual components would be subjected to in the aircraft. They are placed, when possible, so as to take advantage of the radiation gradient which occurs across the panel as a result of the varying distances from the reactor and the shielding of some parts by others.

After final adjustments are made and proper as-

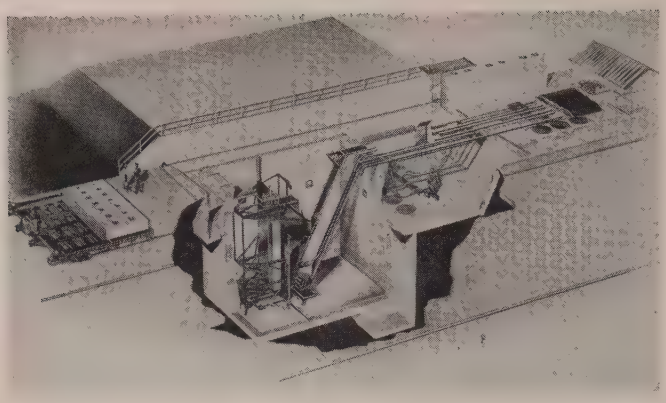


Fig. 10 — Shuttle irradiation system

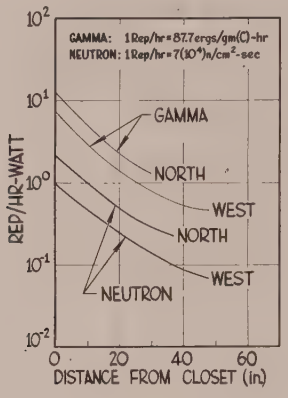


Fig. 11 — Centerline dose rates

sembly operation is established, the assembly is operated and sufficient data are accumulated to establish the operational characteristics of the assemblies. These data provide the criteria for checking satisfactory performance of the assemblies during and after irradiation.

Irradiation — The facility for irradiating the assemblies is shown in Fig. 10. This irradiation system is a modified swimming-pool reactor facility with provisions for remotely locating tests specimens next to the reactor. The pool has been divided into two parts by a dam. The one-third of the pool to the south of the dam is filled with water and provides a reactor handling tank. The north two-thirds of the pool is a dry area into which the test specimens are transported by the shuttle system for irradiation. The dam is approximately 10 ft from the south wall of the pool. Projecting from the center portion of this dam into the north area is a rectangular closet. The reactor is moved into this closet (directly accessible from the wet side) when irradiations are to be made, providing three sides of the reactor core for the placement of test specimens.

The test specimens are transported on dollies riding on a track system which leads to each of the three positions next to the reactor. The tracks extend from the north ramp area to the top of the pool wall and from there, angle down to the irradiation positions.

The system is designed so that test panels on the east and west faces, having dimensions 36×36 in. in a plane parallel to the reactor closet, will be geometrically centered relative to the reactor. On the north face, this is true for a panel 32 in. wide and 36 in. high. The east and west faces were designed for panels 60 in. deep, and the north face for panels 36 in. deep. Test-panel weights are limited to 2000 lb.

Fig. 11 provides some idea of how radiation intensities drop off across a panel as a function of distance from the reactor. The intensities shown are calculated values and only approximate; however, experimental measurements have shown them to be fairly accurate. It can be seen that radiation intensities can vary by a factor of 15 or more across a panel in the east or west positions because of distance from the reactor alone.

After installation in the radiation position, the assemblies are operated prior to reactor start-up. During this period, "baseline" data are accumulated. Any interferences or peculiarities induced by the test configuration are detected at this time and, when possible, steps are taken to eliminate or minimize any discrepancies as they become apparent.

Test conditions during irradiation are selected which, it is believed, will result in conservative estimates of radiation damage. These test conditions are presently based on a limited knowledge of nuclear and nonnuclear environmental effects and interactions. As additional information is provided by the environmental studies efforts, irradiation environmental conditions will be modified as required.

In establishing the length of irradiation, consideration must be given to the possibilities of rate effects. It is not practical to irradiate the assemblies for a period equivalent to their aircraft service-life requirements. We have selected irradiation periods for these tests of 50–100 hr. This results in rates well above those expected in an aircraft, but it is believed that test results will most often be conservative; that is, if rate effects exist, increased rates will result in increased damage.

Having selected an irradiation time and established the dose requirements of the components, the reactor power level (or radiation rate incident on the panel) remains to be established. The physical separation of subsystem components in the aircraft obviously cannot be maintained in assembling the subsystem for irradiation. Therefore, within a single assembly, the neutron-to-gamma ratio and the radiation rate requirements sometimes vary greatly from component to component. As previously stated, in an effort to adjust to these varying requirements the components are arranged in the test panel to take optimum advantage of separation distance from the reactor and shielding by other components. This, however, does not provide a complete solution. In selecting a reactor power level for the irradiation, the component having the highest radiation dose requirement is used. This requirement is specified by both the neutron and gamma dose necessary. The neutron-to-gamma ratio at the irradiation site is near constant throughout. It is not practical to attempt to modify this ratio. However, the reactor power level is selected such that the component will receive either the gamma or neutron dose specified and an overdose of the other (neutron or gamma, as the case may be). This practice results in some components receiving excess doses of neutron and/or gamma radiation. However, conservative estimates of damage in the aircraft application should be obtained.

During the irradiation interval pertinent data are

accumulated at specified intervals and studied for the detection of changes. During periods when changes are occurring the frequency of data taking is increased.

After reactor shutdown, data are also taken before removal of the assemblies from the irradiation area. This is done to observe any changes or recuperative processes which might occur following irradiation. After removal from the pool it is possible to continue operation for extended periods in the handling cave without disconnecting the assemblies.

Post-irradiation testing is conducted under the same conditions as the pre-irradiation tests. Each assembly is disassembled and component tests are conducted and its constituent materials and parts are inspected for evidence of radiation damage and impending failure.

Future Tests

This approach alone cannot provide all the solutions to the development of radiation-resistant subsystems (Fig. 12). However, based upon an analysis of the results of these first tests, some solutions to the problem will be obtained and at the same time a more exact definition of requirements for material development and redesign will be provided. As a second step, assemblies can be constructed and irradiated to determine the feasibility of the indicated design modifications. In the event the first two approaches do not yield completely satisfactory solutions, the relocation and/or shielding of some critical items might be considered — although this will probably not often, if ever, be required.

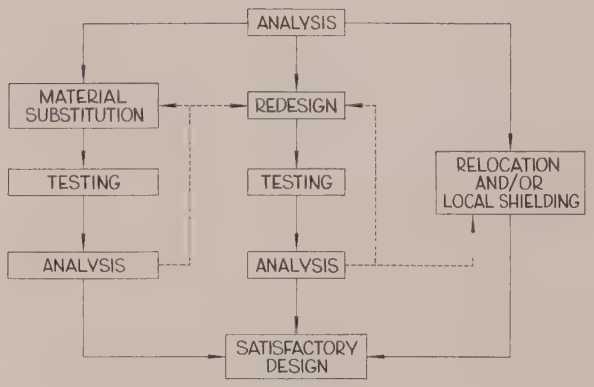


Fig. 12 — Subsystem development

Concluding Remarks

Although there is much which remains to be learned about this new environment, existing technology is such that it is presently possible and desirable to proceed with the building of a nuclear-powered aircraft.

It is only through the experience gained in building and flying this aircraft that the full potentialities of nuclear-powered flight can be realized. The technological advancements which will result in the accomplishment of this task can provide the foundation for not only more advanced earth-bound craft, but for opening the door to nuclear-powered space travel.

References

1. "Convair Radiation Effects Testing System," by J. W. Allen. Convair–Fort Worth Report FZM–1153. Paper presented at Radiation Effects Symposium, Atlanta, October, 1958.

2. "Effects of Nuclear Reactor Radiation on Operation of Aircraft Type Gyros," by M. L. Baker and E. P. Gammon. Convair–Fort Worth Report FZM–1025. Paper presented at Symposium on Radiation Effects, Columbus, Ohio, October, 1957.

3. "Weakest Link Method Helps System Design for Radiation Environments," by H. R. Dvorak. *Space Aeronautics,* Vol. 20, December, 1958, p. 20. (Contains extensive bibliography.)

4. "Effects of Reactor Radiation on B–36 Aircraft Systems." Convair–Fort Worth Report MR–N–167 (NARF–57–41T, August, 1957), Addenda I–V (December, 1957), Addenda VI, VII (August, 1957).

5. "Effects of Reactor Radiation on the Electrical Properties of Electronic Components," Parts 1–5. Convair–Fort Worth Reports MR–N–116 (NARF–56–6T, December, 1956), MR–N–173 (NARF–57–52T, November, 1957), MR–N–178 (NARF–58–2T, January, 1958), MR–N–185 (NARF–58–6T, January 1958), MR–N–216 (NARF–58–35T).

6. "Elastomers for Use in Radiation Fields," by Robert Harrington. *Rubber Age,* Vol. 81, September, 1957, p. 971.

7. "Effect of Nuclear Radiation from GTR upon Closed Circuit Television Equipment," by D. L. Jacobs. Convair–Fort Worth Report FZM–925, June, 1957.

8. "Effects of Reactor Radiation on Electrical Properties of Capacitors," by W. D. McMillan. Convair–Fort Worth Report FZM–1010, October, 1957.

9. "Radiation Effects Problems in Electronic Design," by P. S. Miglicco and D. M. Newell. Convair–Fort Worth Report FZM–1815, June, 1959.

10. "Radiation Effects on Electronic Components," by P. S. Miglicco. *Electric Equipment,* Vol. 5, July, 1957, pp. 26–31.

11. "Nuclear Radiation Effects Research on Aircraft Materials." Convair–Fort Worth Report MR–N–239 (NARF–59–26T, to be published).

12. "Procedures for System Panels Test No. 2." Convair–Fort Worth Report MR–N–122 (NARF–56–27T, September, 1956), Addenda I–IX (September, 1956), Addendum X (January, 1957).

13. "Radiation Effects — Methods and Data." Convair–Fort Worth Report FZK–9–134 (NARF–58–43T, October, 1958).

14. "Results for System Panels Test No. 2." Convair–Fort Worth Report FZK–9–126 (NARF–58–1T, January, 1958), Addenda I, II (January, 1958), Addenda III, IV (December, 1958), Addendum V (September, 1958), Addendum VI (June, 1958), Addendum VII (December, 1958), Addendum VIII (May, 1958), Addendum IX (December, 1958).

15. "Transient Irradiation Response of Semiconductors," by A. B. Spears. Convair–Fort Worth Report FZM–1722, January, 1959.

16. "Some Effects of Pulsed Neutron Radiation on Electrical Components." IBM Corp. Radiation Research Group, October, 1958.

17. "Effects of Nuclear Radiation on Electronic Components." Admiral Corp. Scientific Reports 1 through 10, October, 1955 through June, 1958.

18. "Effects of Nuclear Radiation on Seals, Gaskets, and Sealants." Radiation Effects Information Center, Technical Memorandum No. 8, Battelle Memorial Institute, November, 1958.

THIS IS a progress report on the AMA-TTMA program to develop a brake horsepower absorption rating system. The authors outline the need for service brake rating and the application of rating methods to determine the braking ability of commercial vehicles.

The first part of the paper describes the problem which faces the motor truck industry in the establishment of a rating formula. The second part presents one possible method. Both authors agree that many problems are still to be solved.*

Horsepower for commercial

John H. King, Automobile Manufacturers Assoc.

Alex F. Stamm, Chrysler Corp.

Part I by John H. King

COMMERCIAL VEHICLE and vehicle combination gross loads have been rising in recent years. Greater emphasis has been placed on ability of such vehicles and combinations to maintain adequate speeds in mixed traffic on major highways and expressways.

The motor truck and trailer manufacturing industries hope, and expect, that these trends in higher gross loads will continue in the years ahead. We have, in fact, recently proposed to the American Association of State Highway Officials that substantially increased vehicle loadings be permitted in coming years as the nation's state-federal highway modernization program moves forward.

These trends have resulted, however, in a growing dissatisfaction by many state vehicle regulatory agencies with their current methods for preventing commercial vehicle overloads, with particular reference to engine power and the problem of brake fade in mountainous terrain. These regulatory officials have come to the industry for recommendations as to how to deal with these problems.

When the industry proposed that manufacturers' gross vehicle and gross combination weight ratings (mgvw and mgcw) be used as a method for control of overloads — with such tolerance factors as individual states might determine from past experience to be within acceptable margins for safety and adequate performance — the objection was raised that these ratings are established on a basis of engineering judgment and experience by each vehicle manufacturer. Fears were expressed that such ratings might be excessively upgraded unless they were established by known criteria which could be checked by independent tests if the states desired.

Although the industry felt these fears were unfounded (because warranty considerations and the manufacturers' own reputations would be affected if ratings were not based on sound engineering elements), the states have been reluctant to use them unless ratings are set on a basis of known engineering factors that can be validated by objective tests.

As a result, other types of controls were sought — and in at least one important state they were actually adopted several years ago. Because such controls directly affect vehicle engineering and design features, they can act to stifle future progress in vehicle safety engineering and performance. In addition, any further spread of such controls invites economic chaos in truck manufacturing, with serious penalties for manufacturer and operator alike.

The AMA Motor Truck Committee, therefore, decided several years ago that the future expansion of motor truck transportation requires that certain minimum engineering standards be established for mgvw and mgcw ratings, and that these minimum standards be set forth in a published code. By legal reference to such a code, public legislative and regulatory agencies then can define precisely what is required in manufacturers' certifications of gross vehicle and gross combination weight ratings, while at the same time avoiding detailed regulations on engineering design elements.

We have completed three of four projected elements in this proposed code. They cover a net engine horsepower requirement related to gross weight, a listing of certain vehicle components and general rules for their selection as related to gross weight ratings, and an explanatory statement which points out that gvw and gcw ratings are necessarily conservative and sometimes can be exceeded with safety under favorable conditions of speed, terrain, and road surface.

In other words, the industry is deliberately seeking to encourage individual states and municipalities to establish their own tolerance factors, based on their own experience, in using gvw and gcw ratings as a basis for vehicle gross weight controls. We intend merely to establish uniform ground rules for the manufacturers' ratings, to publish these ground rules, and to seek to an appropriate degree to have at least some of the rules adopted as Recommended Practices of the Society of Automotive Engineers.

We know of no other way to approach the problem of controlling a small minority of vehicle operators who load vehicles far in excess of their design capacity. We recognize that there are problems here, because under favorable operating conditions the ratings can be exceeded to a modest degree without prejudice to vehicle safety or satisfactory service life. Yet it is not possible to spell out by some fixed

* Paper presented at SAE Annual Meeting, Detroit, Jan. 13, 1960.

Rating vehicle brakes

formula the varying degrees of flexibility that might be justified.

The final element of our rating approach, which is as yet incomplete, involves a rating method for brake horsepower absorption, and a requirement on minimum brake horsepower absorption capacity per unit of gross vehicle or gross combination weight. This is the subject of this discussion.

The assignment to develop this brake horsepower absorption rating system was given by the AMA Motor Truck Committee to the Joint Technical Committee on Combination Vehicle Brakes, which represents both the AMA Motor Truck Technical Committee and the Engineering Committee of the Truck-Trailer Manufacturers Association. The Joint Committee also has, as continuing consultants, representatives of 13 companies which supply components of braking systems.

It has been a difficult undertaking, because many installation variables need to be taken into account in rating actual performance of brakes on a given vehicle during actual highway operation.

Initially, the Committee undertook the drag-test approach, the object being to maintain a constant measured braking force with the use of a towing dynamometer. Problems arose which seemed insurmountable, due to the low level of braking input necessary to extend the test period to a desirable time level. Reliable measurements and accurate braking control could not be achieved.

Tests of steady braking input may become feasible on a laboratory dynamometer, but the question of correlating such tests with actual highway conditions still remains.

After the continuous-drag tests were abandoned several years ago, the Committee decided upon a 10-min test, involving a uniform cycle of brake snubbing from 40 to 20 mph, with the objective of adjusting the cycle to attain a 10-ft deceleration rate on all snubs except the final snub, which was to involve a near-fade condition under which the deceleration rate would decline to 8 ft.

It was specified that line pressure in braking at no time during the test should exceed double the lowest pressure. Except for this requirement, and the one concerning deceleration rates, no other ground rules were found necessary. (This is a slight oversimplification, because certain adjustments were permitted in test vehicle weight and deceleration rates when necessary to bring brakes to a near-fade condition. But basically the test rules were quite simple.)

Under an equation developed by the Committee, which took parasitic braking into account, an equivalent continuous horsepower absorption rating could be computed for the 10-min cycle.

For the most part, initial testing by a number of vehicle manufacturers produced reliable ratings that could be duplicated rather consistently in repeated tests. The Committee became convinced that a practical testing method had been developed. Therefore, it undertook to study the effects of installation variables, with particular stress on different types of brake linings. The objective was to determine to what degree these variables changed the ratings, with the ultimate aim of developing points of reference that would permit transfer of the entire road-testing program to laboratory dynamometers after correlation factors had been fixed.

Here, unfortunately, we have met new obstacles. It appears that certain linings increase in efficiency with each successive snubbing test, until they reach a point of destruction which often is difficult to detect in the road testing. They stubbornly refuse to fade to the required 8-ft deceleration rate. While these same linings may have other characteristics that make them undesirable for any but specialized commercial use, they are capable of achieving excellent ratings on the horsepower absorption tests. It appears, then, that the Committee will need to amend its ground rules. We believe the obstacle is going to be overcome; but how it will be done is not yet clear.

We have been encouraged by the fact that motor vehicle manufacturers and vehicle regulatory agencies in many overseas nations have given an enthusiastic reception to reports on the work of our AMA–TTMA Brake Committee in this horsepower absorption rating approach.

British vehicle manufacturers conducted intensive tests of our approach and concluded that it is a practical answer to the problem of rating antifade characteristics of brakes. The British approach involves a slight modification of our concepts. They qualify a braking system if, after a specified snubbing test cycle, a measured stop can be made which shows that the hot brakes are at least a specified per cent as efficient as were the brakes prior to the test cycle. While this does not produce an actual horsepower absorption rating, for enforcement of legal gross load limits it seems a satisfactory approach.

Part II by Alex F. Stamm

STOPPING ABILITY as an index to brake performance is well established, but a good braking system also has other characteristics which must be evaluated and fall within acceptable limits.

Some interesting relations can be developed by considering energy relations involved in the stopping of a vehicle as compared to holding a steady speed down a given grade. Two important input conditions can be recognized, the transient phase of stopping or bringing about a decrease in speed, and the steady or sustained state where the vehicle speed remains substantially unchanged but a more or less constant input into the retarding mechanism is necessary to balance the effect of gravity. The

transient condition is characterized by very high input rates involving comparatively small amounts of energy. On a downgrade contrary conditions prevail, input rate being quite nominal but the total energy transferred can be large. As a specific example, any vehicle can be expected to descend a 4% grade, say 5 miles long, at a speed of 30 mph. The time of such a descent would be 10 min. To transfer the same total energy on the level would require the making of 35 stops from 30 mph in 10 min. (For the present, effect of parasitic braking is neglected.) If it is assumed that the stops are made at the Uniform Vehicle Code standard required ability of 14 fps^2 deceleration rate, it would theoretically take 3.15 sec per stop, leaving about 14 sec to get back up to speed. Each 128 lb of vehicle would need 1 hp just for acceleration. Taking account of other horsepower demands brings up the net horsepower requirements, leading to an estimate of at least 479 net engine horsepower needed for a 50,000-lb vehicle. Light vehicles generally do have sufficient engine power to permit this cycle of operation, but many heavy combinations could not be brought up to 30 mph in a sufficiently short time to match the requirements of the downhill run.

Every truck must be capable of reasonable downhill performance and the example selected for discussion here can be commonplace in some localities. It will be recognized that a series of rapid stops or snubs can, in some respects, be taken as the analog of the hill. This procedure may become the basis for a second brake capability parameter if it is possible to agree upon and recognize an acceptable rating cycle.

Further examination of energy and power relations will give some conception of relative magnitudes essential to understand and differentiate the two braking conditions which have been defined. An infinite number of combinations of speed, load, grade, and distance are possible. It is also worth remembering that a high deceleration rate is not often used in normal stopping, with the consequence that stopping brake horsepower is generally much lower than the values shown in Table 1. It is emphasized that the example was arbitrarily selected and does not necessarily represent any specific vehicle or operation. Summarizing the assumed conditions more completely, gross weight used in the computations is 50,000 lb, speed is taken at 20, 30, and 40 mph, grade is constant at 4%, and the time on the grade is kept at 10 min. The energy transfer in stopping is independent of stopping rate, being simply the kinetic energy of the vehicle ($KE = \frac{1}{2} MV^2$). Similarly the potential energy change in going down a hill is the vertical descent times the vehicle weight. Table 1 summarizes the three cases for the transient condition, giving initial speed, stopping time and distance at 14 fps^2 deceleration, energy dissipated, and rate of energy transfer. A corresponding steady state tabulation is also shown, listing, for three conditions of speed sustained for 10 min down 4% grade, distance, energy dissipated, and rate of dissipation. Since a steady speed is maintained, it is feasible to estimate the dissipation due to tires, chassis friction, and air resistance. Calculations by TR-82[1] are shown as parasitic horsepower, and the net input rate into the brakes and/or retarding system is shown.

A similar analysis by TR-82 can be made to give the net engine horsepower needed to climb the same hill. It can easily be shown that it takes considerably more engine power to get the vehicle up the hill than to come down on the brakes at the same speed. The important conclusion is that the brakes are likely to work much harder in going downhill than in the normal stopping requirements on the level road.

A vehicle moving at uniform speed increases its distance, from any arbitrary reference point, uniformly with time as shown in the upper curve of Fig. 1. In general terms, the motion is defined by $x = V_1 t$ and x, the distance, is given in terms of time t on a time basis. Stopping the vehicle at a uniform rate of negative acceleration can also be written on a time basis. Now the lower curve in Fig. 1 shows what happens if the initial speed is V_1 and the stopping rate is $-a$ fps^2. The vehicle will have zero velocity after a certain time and it will travel a smaller distance while decelerating than it would have in the same time without deceleration. Slowing down at a uniform rate is given on a time base by the equation:

$$x = V_1 t - \frac{at^2}{2} \qquad (1)$$

x is maximum when $\frac{dx}{dt} = 0 = V_1 - at$ or when $t = \frac{V}{a}$

Therefore, $x_{max} = V_1^2/2a$

Now consider that the vehicle does not stop completely, but decelerates uniformly from the initial velocity V_1 to a lower speed speed and then continues at a new lower velocity. Time and distance continue to increase indefinitely, but, as the middle curve in Fig. 1 shows, the lower portion is nonlinear until deceleration ceases and then the rate of distance increase becomes constant at a lower speed. The function is discontinuous, being the sum of the two earlier considered cases taken consecutively in proper order and over the correct intervals. It is this difference, that the two equations are not added simultaneously, which characterizes the function as discontinuous. Nevertheless, a useful equation of motion can be written in which time is divided into two distinct parts:

$$x = \left[V_1 t - \frac{at^2}{2} \right]_0^{T_1} + [V_1 - aT_1] \cdot [t - T_1] \qquad (2)$$

Acceleration has a minus sign because speed is decreased and T_1 is the time during which the acceleration a is acting. The first quantity applies only while speed is changing, with t never becoming greater than T_1. The second quantity is the new velocity and the third term is the time at this new speed. Time has only positive values; negative values of the final quantity are neglected, and T_1 can equal but not exceed $V_1 a$, which makes the right term equal zero. Any desired schedule of vehicle motion can be expressed in a similar manner.

Before going to the next step, it may be well to point out once again that steady state conditions have some limitations in their application to braking system investigations when applied to a vehicle for the purpose of brake rating. A constant deceleration rate cannot be maintained for a long time, since this would require either that the rate ap-

[1] "Truck Ability Preduction Procedure." SAE Special Publication TR-82, fifth edition, 1960.

proach zero or that the initial velocity be infinite. Longer steady state conditions can be maintained on a hill, but this is often not convenient. An equivalent condition independent of special terrain is much to be desired. Some exploratory work has been done by drag testing, the object being to maintain a constant measured braking force with the use of a towing dynamometer. Difficulty was encountered due to the low level input rate necessary to extend the rating time to a desirable interval. Reliable measurements and accurate braking control could not be achieved. The prospects for steady force testing may be more attractive on a dynamometer when limited to the brake alone, but the question of correlating such tests to the actual installed conditions still remains. A simple way of obtaining useful results directly on the vehicle would be the best answer.

A recurrent cycle, consisting of negative and positive accelerations, is better suited to this purpose. Such a cycle will, in general, be quite unsymmetrical, with the retarding portion being a relatively small part of the period. Its superiority lies in the fact that the vehicle itself supplies the input, and frequency as well as intensity can be adjusted as needed over a very wide range. Such a snubbing cycle is shown on a time base in Fig. 2. Distance is expressed by the discontinuous two-part equation:

$$x = \left[V_1 t - \frac{at^2}{2}\right]_0^{\frac{V_1 - V_2}{a}} + \left[V_2 t + \frac{1}{2}\left(\frac{V_1 - V_2}{T - \frac{V_1 - V_2}{a}}\right) \cdot t^2\right]_{\frac{V_1 - V_2}{a}}^{T} \quad (3)$$

The first term represents the snubbing part of the cycle between speeds V_1 and V_2 at $-a$ acceleration. In the second term the speed changes from V_2 back up to V_1 at a rate so the total period of one complete cycle will be T. Any number of such complete cycles of a snubbing test can be repeated. Distance will increase in a nonlinear manner, but there is a ray containing all points at which the sign of the acceleration changes; consequently, at the intersections distance has increased by an amount given by the average of the maximum and minimum velocities of the cycle times the elapsed time.

Energy is delivered into the brakes during the snubbing part of the cycle only; during the time interval t_b. The distance equation is, once again:

$$x = V_1 t - \left(\frac{at^2}{2}\right) \quad (4)$$

and: $V = \frac{dx}{dt} = V_1 - at$ (a linear equation)

A sawtooth velocity-time wave as shown in Fig. 3 results from differentiating the distance curve of Fig. 2. The snubbing portion is linear. If acceleration is at a uniform rate, this part will also be linear but, as far as the input into the brakes is concerned, the path in going from A to B is unimportant as long as the time of the period, T is maintained. Points A and B are sufficiently defined as point functions for braking input considerations. The shape of the acceleration cycle may, however, have a secondary

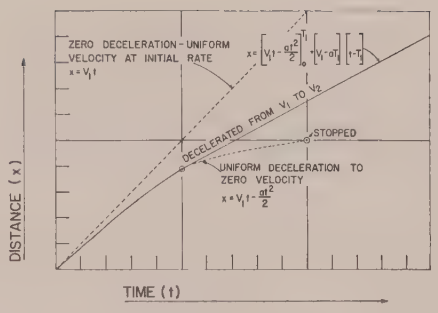


Fig. 1 — Effect of various deceleration rates on vehicle

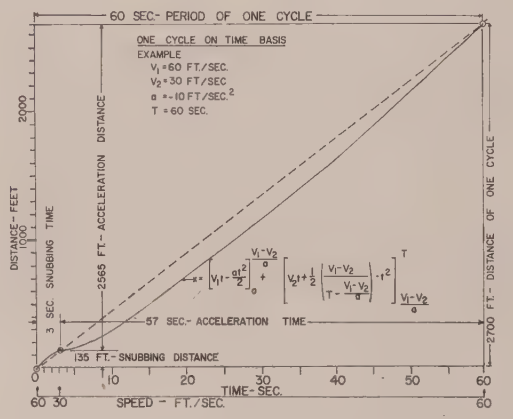


Fig. 2 — Snubbing cycle

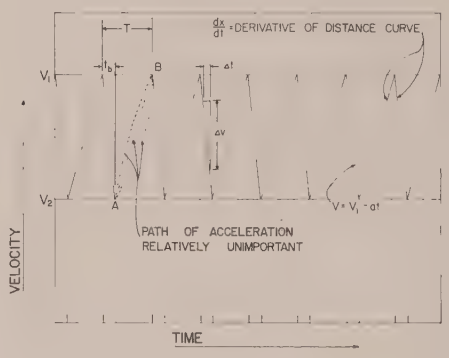


Fig. 3 — Velocity-time wave

Table 1 — Stopping Brake Horsepower

Transient Condition
stopping 50,000 lb at 14 fps²

Initial Speed		Stopping Time, sec	Stopping Distance, ft	Energy Dissipated		Total Stopping Hp	
Mph	Ft/sec			Ft-lb	Btu	Peak	Average
20	29.3	2.10	30.6	$\frac{2}{3} \times 10^6$	858	1160	580
30	44.0	3.15	69.2	$1\frac{1}{2} \times 10^6$	1930	1740	870
40	58.7	4.20	123.0	$2\frac{2}{3} \times 10^6$	3435	2320	1160

Steady State Condition
descending 4% grade in 10 min with 50,000 lb gvw

Sustained Speed		Distance		Energy Dissipated		Horsepower		
Mph	Ft/sec	Miles	Feet	Ft-lb	Btu	Total	Parasitic	Net
20	29.3	$3\frac{1}{3}$	17600	35.2×10^6	45200	106.7	64.6	42.1
30	44.0	5	26400	52.8×10^6	67800	160.0	89.4	70.6
40	58.7	$6\frac{2}{3}$	35200	70.4×10^6	90400	213.5	124.5	89.0

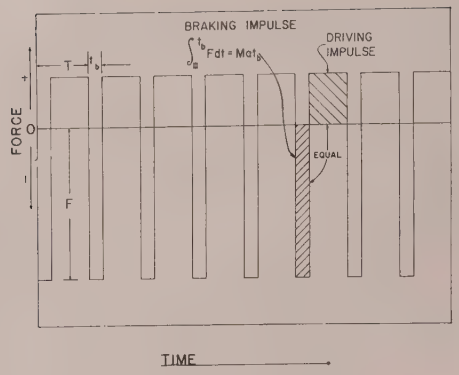


Fig. 4 — Force cycle

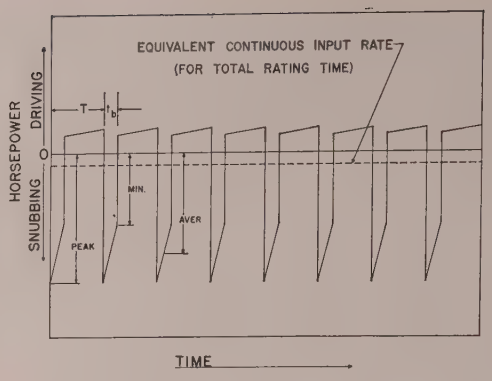


Fig. 5 — Periodic horsepower function

effect on brake capacity because cooling is a function of vehicle velocity. Consequently, it appears desirable to maintain this part of the cycle as nearly linear as convenient. However, during a practical series of rating cycles the input into the brakes is relatively large and the total time (10 min) rather short in relation to the time necessary to cause appreciable temperature drop of the brake. Finally, the time of a cycle is in the general order of ½ min, all this substantiating the contention that such deviation from linearity as will be encountered in practice during acceleration is, in fact, of little importance.

By virtue of vehicle mass, a force must be applied to change the velocity. The force will be a constant as long as acceleration is maintained unchanged. In a recurrent cycle, assuming 100% efficiency the impulse imparted by the total braking effort, equals the engine supplied impulse. The force cycle is the derivative of the velocity wave times the system mass. It is the unsymmetrical square wave shown in Fig. 4. Of course, in all these schematic representations true proportions have been ignored to emphasize the principle. The areas above and below the zero axis in Fig. 4 are equal but the accelerating time is in the order of 10 times the snubbing interval, indicating some distortion in the time scale shown. The upper part of the square wave will, in general, not be linear, but again, as long as the areas balance, the cycle will be maintained. Work is represented by the integral:

$$W = \int F\, V\, dt$$

Considering the braking part of the cycle and substituting Eq. 4 for the velocity:

$$\text{Work per cycle} = \int F\, V_1\, dt - \int F\, a\, t\, dt$$

$$= \left| F\, V_1\, t - \left(\frac{F\, a\, t^2}{2}\right)\right|_o^{t_b}$$

$$= F\, V_1\, t_b - \left(\frac{F\, a\, t_b^2}{2}\right)$$

but: $$t_b = \frac{V_1 - V_2}{a},$$

hence:

$$\text{Work per cycle} = F\left[\frac{V_1\,(V_1 - V_2)}{a}\right] - \frac{a}{2}\cdot\frac{(V_1 - V_2)^2}{a^2}\Bigg]$$

$$= \frac{F}{2a}\left[2V_1^2 - 2V_1V_2 - V_1^2 + 2V_1V_2 - V_1^2\right]$$

$$= \frac{F}{2a}\left[V_1^2 - V_2^2\right] = M\left[\frac{V_1^2 - V_2^2}{2}\right] \quad (5)$$

Horsepower is the rate of doing work. Braking force, equal to Ma, acts over the velocity range from V_1 to V_2. At any instant while braking the velocity is $V_1 - at$. This gives the horsepower equation as:

$$hp = \frac{Ma}{550}(V_1 - at) \quad (6)$$

A similar relation exists while accelerating, thus the periodic horsepower function appears as shown in Fig. 5. During the braking part of the cycle the power is given by:

$$hp_{\text{peak}} = \frac{Ma\,(V_1)}{550} \qquad hp_{\text{min}} = \frac{Ma\,(V_2)}{550}$$

$$hp_{\text{average}} = \frac{Ma(V_1 + V^2)}{1100}$$

The horsepower input to the braking system is seen to consist of a series of sharp negative going pulses corresponding to the intermittent retarding pulses shown in Fig. 4. During this time the brake is absorbing energy. Dissipation of the braking energy commences at the moment the first pulse is initiated and increases in a complex manner throughout the entire test time and continues at a decreasing rate for a long time afterwards until final thermal equilibrium is reestablished. Only a part of the energy imparted by each pulse is lost before the next pulse occurs. Storage in the system accounts for the remainder, causing temperature and rate of dissipation to become greater. At the end of the last pulse, the amount of energy stored and the rate of dissipation, as well as temperature, are all at a maximum. It is, therefore, logical to express the input into the braking system in terms of an equivalent continuous rate sustained for the total time of the test. This must not be confused with dissipation and it should be clear that input and output are equal in terms of energy but not when expressed as a rate.

Dividing the work input per cycle given by Eq. 5 by the total period of one complete cycle and converting this to horsepower gives:

Equivalent Continuous Braking

$$\text{Input Horsepower} = \frac{M}{1100T}(V_1^2 - V_2^2) \quad (7)$$

Most of the energy of the snubbing pulse can be channeled into the brakes but a relatively small fraction will inevitably be dissipated by parasitic losses, including tire and air resistance and chassis friction. A partial correction can easily be made based upon the assumption that parasitic losses result in a constant force during the braking cycle,

that is, a fixed percentage of deceleration rate used is chargeable to losses.

Drift time, defined as the time for the vehicle to decelerate between the speed limits used in the snubbing cycle with all conditions unchanged, except that the brakes are not applied, can be converted into an equivalent sustained parasitic retarding force.

Since:

$$F_o = Ma_o \text{ and } a_o = \frac{V_1 - V_2}{t_o}$$

$$F_o = \frac{M(V_1 - V_2)}{t_o} \quad (8)$$

Subscript $_o$ refers to parasitic draft time and force.

Net force developed by the brakes during a snub cycle must then be:

$$F_{\text{net}} = Ma - \frac{M(V_1 - V_2)}{t_o}$$

Work, the product of force and distance or force times time times average velocity, per snubbing cycle becomes:

$$W_{\text{net}} = \underbrace{\left[Ma - \frac{M(V_1 - V_2)}{t_o}\right]}_{\text{Force}} \times \underbrace{\left[\frac{V_1 - V_2}{a}\right]}_{\text{Time}} \times \underbrace{\left[\frac{V_1 + V_2}{2}\right]}_{\text{Average Speed}}$$

(Time × Average Speed = Distance)

Corrected equivalent continuous braking input horsepower is $\frac{W \text{ net}}{550T}$, or:

$$\text{hp}_{\text{net eq.}} = \frac{Ma}{550T}\left[\frac{V_1^2 - V_2^2}{2a}\right] - \frac{M}{2 \times 550T}\left[\frac{V_1^2 - V_2^2}{t_o}\right]\left[\frac{V_1 - V_2}{a}\right]$$

Substituting: $t_b = \frac{V_1 - V_2}{a}$

$$\text{hp}_{\text{net equiv.}} = M\,\frac{(V_1^2 - V_2^2)}{1100T}\left(1 - \frac{t_b}{t_o}\right) \quad (9)$$

This is the same as Eq. 7 multiplied by a correction factor.

Eq. 9 is a logical basis for defining braking system equivalent steady input horsepower for a specified time. The rating interval is divided into a number of equal periods during each of which the speed changes from V_1 to V_2 under the influence of the brakes and back from V_2 to V_1 by the engine. If a reasonably large deceleration rate is used, the correction factor approaches unity and the brake horsepower becomes almost independent of deceleration rate. Also if V_2 is taken as $\frac{1}{2}\ V_1$ the value of $(V_1^2 - V_2^2)$ is $\frac{3}{4}\ V_1^2$, indicating that the maintaining of exact values of V_2 is of secondary importance. In actual testing it is convenient to avoid bringing V_2 to zero, as it is much easier to accelerate from, say, 20 to 40 mph than from zero to the same speed in the necessary time. Horsepower thus becomes a function of total mass controlled by the brakes, the speed range used, and the number of snubs employed in the rating time. It is this last variable, the number of snubs, that is most importance in making a rating test. That no standard rating procedure has been agreed upon so far can be attributed to the difficulty in defining the condition of the braking system determining the end of the test. The rating is intended to be nondestructive except for normal wear; therefore, a method of gaging just when the brake has been subjected to the maximum input without permanent damage must be developed.

DISCUSSION

Snub Test Presents Several Problems

— D. P. Fisher
Chevrolet Motor Division, GMC

HIS PAPER points out the need for a reliable method of rating motor vehicle brakes, and touches briefly on the background of testing toward establishment of such a method of brake rating.

Reference was made to a continuous drag type of test, and the difficulties encountered. To elaborate on this for just a moment — initial tests were of only 3-min duration at a high rate of energy input, but it was found that the brake drum and other components could absorb the heat generated during this short test period. When the test time was extended to 10 or 15 min to increase the energy input beyond the brake and component absorption capacity, the rate of energy input had to be reduced to such a low level that the *range* of *variation* in drawbar pull was greater than the *nominal*. Factors causing the variation were towing vehicle driver error, wind velocity and direction, and condition of the test road with respect to gradient and waviness.

The problem was further complicated by the degree of friction in the brake mechanism; for example, one manufacturer, in conducting tests on the air brakes on the rear of one of their larger models, found it necessary to use a Type 24 chamber on one side, and a Type 30 chamber on the other. A balance of the friction and spring loads between the two brake assemblies could not otherwise be obtained at the very low chamber air pressure used during the test. Since similar problems were encountered with hydraulic brakes and, further, since this type of test required two vehicles, three operators, sensitive drawbar equipment which was easily damaged, and efficient communication between the towing and towed vehicles, this type of test was abandoned in favor of a repeated snub type of test.

Problems were immediately encountered in the snub test also, but these were solved satisfactorily as tests progressed, except for two which are yet unresolved:

1. A means of determining when the brake lining is not only burnished for full drum contact, but also properly conditioned for maximum effectiveness.

2. A means of determining when the maximum brake rating has been obtained without having permanently damaged the brake lining.

Up to the present time, the first of these two problems appears to be the more complex, since some linings require more energy input than others to reach maximum effectiveness. A definite means has not yet been agreed upon for determining when a set of linings are properly conditioned for giving a maximum brake rating, short of a complete series of tests to determine the consistency of test results.

The second problem is currently being investigated from the standpoint of a recovery test following the rating test proper, which is basically a fade test. Baseline performance is established prior to running the rating test by making three "snubs" from 30 to 20 mph at a deceleration rate of 10 fps^2 at ½-mile intervals, recording the application line pressure. Following the rating test, a minimum of six snubs are made 30 to 20 mph at a deceleration rate of 10 fps^2 at 2-mile intervals, recording the application line pressure. The allowable variation, plus or minus, over the baseline pressure has not as yet been agreed upon, due to a lack of data on characteristics of

various brake linings.

Perhaps it will have occurred to some who are directly involved in brake testing that the brake rating can be directly affected by variation in cycle speeds or deceleration during the rating test. This is true, and since it is unreasonable to expect any driver to obtain exactly 10 fps^2 deceleration from exactly 40 mph to exactly 20 mph on each and every brake application, it is important to know which values have the greatest effect on test results.

Fig. A shows the effect on horsepower input, per cycle, of any variation in the deceleration rate for that cycle. Four curves are shown representing cycle times of 20, 30, 40, and 50 sec. In the rating test range of 10 fps^2, the slope of the curves is very flat; therefore, an error of even 10%, such as 9 or 11 fps^2 instead of 10, has only a very slight effect on horsepower input, in fact, less than 1%.

Fig. B shows the effect of variations in either the initial or final cycle velocities. Two groups of curves are shown, each again representing cycle times of 20, 30, 40, and 50 sec, plotted to show velocities of 20 and 40 mph, ±8 mph. The group on the left shows the effect of starting the cycle at exactly 40 mph, with a variation in the final velocity. Note that with a 20-sec cycle, a final velocity of 19 or 21 instead of 20 mph would affect the horsepower input for that cycle to the extent of 2.9%. The group of curves on the right shows the effect of variation in the initial velocity with the final velocity exactly 20 mph. Here, a variation of 1 mph in initial velocity affects the horsepower input for that cycle to the extent of 6.8% with a cycle time of 20 sec, or 7.9% with a cycle time of 50 sec.

Variation in the cycle *time* during a specific test is also very important, but this factor is more easily controlled than, for example, the rate of deceleration. The actual effect of a 1-sec variation on the horsepower input per cycle is, of course, greatest when the time cycle is shortest; for example, should the established cycle time be 20 sec, a 19-sec cycle would increase the horsepower input by 5.3% for that cycle, while a 21-sec cycle would decrease the input by 4.8% for that cycle only.

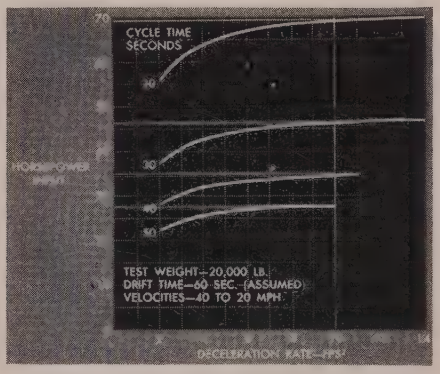


Fig. A — Horsepower input versus deceleration rate, equivalent snub test

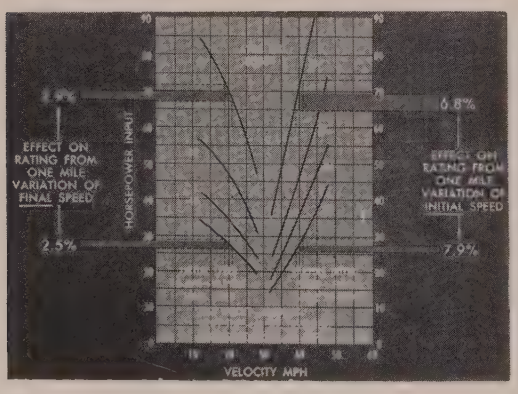


Fig. B — Horsepower input versus velocity, equivalent snub test

It has now been made clear that in the order of their effect on the horsepower input per cycle, the initial cycle velocity is of prime importance, ranging from 6 to 8%; cycle time is second in importance, ranging from 2 to 5%; final cycle velocity is third with a range of 2 to 3%; and the deceleration is fourth — less than ½ of 1% depending on cycle time). Fortunately, the two most important factors — initial cycle velocity and cycle time — are also the most easily controlled. Also, experience has shown that the average driver does not consistently err in the same direction — that is, on the high or low side of a given velocity or rate of deceleration. Therefore, these errors in velocities, rate of deceleration, or cycle time tend to cancel out in a test of 10-min duration, consisting of a number of cycles ranging approximately from 12 to 30.

It is anticipated that eventually the tests for brake rating will be conducted on a brake dynamometer, under laboratory controlled conditions. Before this can be done, it is almost unanimously agreed we must know more about the effect on cooling of brake environment, or installation components such as cast or disc wheels, tube or tubeless tires, drum type, and drum projection beyond the rim or tire. When this data is available, and any required corrective factors are established, then it will be possible to establish reasonably reliable brake horsepower ratings on the brake dynamometer.

Heat Dissipation Remains Problem on Current Brake Design

— S. G. Tilden
The Permafuse Corp.

MR STAMM develops the concept and compares the requirements of brakes as used first to decelerate and second to control. To my knowledge he is the first to point out that retardation of the vehicle requires the intermittant absorption of *kinetic* energy, while control in descending grades requires the continuous absorption of *potential* energy.

In doing the first, that is that of absorbing the kinetic energy of the vehicle intermittently, present brake designs do an excellent job and, — rated on horsepower absorption ability — need no apology. The example, given by Mr. Stamm of a 50,000 gvw vehicle decelerating at 14 fps^2 from 40 mph wherein the peak power absorbed was 2320 hp and the average 1160 hp is proof of this.

However, as potential energy absorption devices in continuous service, they fall very short statistically; to wit the figures in Table 1 of being required to absorb only 89 hp over a period of 10 min in controlling this same vehicle down a 4% grade at 40 mph.

In rating the brake on the basis of its ability to absorb kinetic energy for short periods (that is, bring the vehicle to a stop on a level road), we can depend upon the thermal capacity of the brakes and drums to absorb the heat. But in calling upon this brake to absorb the potential energy of control on hills over an extended period, we must look to heat dissipation—not absorption.

Our present orthodox brakes fall short on heat dissipation ability for the simple reason that we have built no heat dissipation means into the device. A means of heat dissipation is essential if we are to improve brakes as continuous heat absorption devices. In other words, if we are going to rate brakes on their continuous service ability, we're going to have to cool them. Such cooling could be by water, as expounded by at least one brake manufacturer, or could be greatly improved aircooling. After all, in aircraft work, there seems to be no great trouble in continuously aircooling a 3500-hp engine and it would be interesting to turn loose a good aircraft engine designer on the problem of cooling present brakes by controlled airflow and thus improve the horsepower rating of the brakes as potential energy absorption devices.

Perhaps there will be a new switch on the instrument panel and when you start down a hill, you'll "turn on the blowers."

a universal means for Rating Diesel Engines for Deposits and Wear

F. A. Robbins, Koppers Co., Inc.
L. G. Schneider, U. S. Naval Engineering Experiment Station
G. H. Shea, Socony Mobil Oil Co., Inc.

THE NEW CRC Diesel Engine Rating Manual is intended to furnish a universal language for identification of diesel-engine deposits and wear.

Diesel-engine pistons are evaluated for lacquer deposits by utilizing an area demerit basis and color gradations of brown and gray from clean to black. In studying various means for evaluating thickness and texture of deposit in oil systems, it was decided that the scratch gage developed by the CRC Engine Deposit Rating Panel of the CRC-Motor Engine Varnish and Sludge Group was suitable for diesel engines.

A procedure for establishing a volume factor which furnishes a weighted interpretation of the deposit was created.*

IN ORDER TO ADVANCE the science of designing, operating, and maintaining diesel engines, cooperative test work has been and is being conducted mutually by engine builders, petroleum companies, and customer groups such as railroads, truck operators, United States Government, and the like. In carrying out these test activities, a need has become apparent for a common basis for evaluating engine performance with respect to deposits and wear.

Individual companies and governmental agencies have established from time to time engine rating procedures to meet their individual requirements in terms of their experience. This has resulted in a great many different approaches to rating engines for deposits and wear; therefore, there is no basis for comparing test results. We have lacked a common language within this area of activity, and the resultant effect on communication and interpretation has been similar to the effect of a language barrier between nations.

This paper introduces the "CRC Diesel Engine Rating Manual (CRC Manual No. 5)" which has been prepared by the Group on Full-Scale Test Techniques of the Diesel Vehicle Fuel, Lubricant, and Equipment Research Committee of the Coordi-

* Paper presented at SAE Annual Meeting, Detroit, Jan. 12, 1960.

nating Research Council, Inc. This Manual is intended to furnish a universal language for identification of diesel-engine deposits and wear.

History

A review of the experience of test groups functioning within the framework of the CRC has indicated the adverse effects of having no sound or systematized, pre-established approach to engine rating. In conjunction with the program for rating submarine engines for the effects of high sulfur fuel at the United States Naval Experiment Station, a special CRC Subpanel operated effectively to revise a system for rating the engines under consideration for deposits and wear. This experience, coupled with the general recognition within CRC for a need of a common approach to engine rating in the period from 1952–1954, culminated in a decision by the CRC Diesel Group on Full-Scale Test Techniques to set up a special project. At a meeting of the Group in October, 1954, at Cleveland, a panel was created with the following purpose:

> ". . . . to collect the various deposit and wear rating systems in use in industry, indicate the difference between the various systems and, on that basis, make a recommendation for a standard system to be used in CRC."

The membership was comprised of five engineers from petroleum companies, three from engine builders, and three from customer organizations.

The initial meeting of the Rating Systems Panel was held in May, 1955, at Bartlesville, Okla. At the time of the formation meeting, the generally recognized techniques for rating engines were identified and distributed to the members of the Group for study. The following decisions were made as to principles by which the work would be conducted:

1. All engine characteristics which could be defined in finite measurements should be so defined.
2. Those engine conditions which could not be measured would be defined on a demerit basis, with 0 being the new condition, and 10 an arbitrary high level considered most unsatisfactory for the part under consideration.
3. The purpose of each rating is to define a condition, and the interpretation thereof would be the responsibility of the laboratory conducting the test.
4. The relative importance of various ratings would not be incorporated, and no means for combining demerits for whole engines or portions of engines would be established.

At the next meeting of the Panel in the fall, subpanels were created to establish the most logical means for rating the different areas of the engine from the available rating systems in the literature and the experience of the laboratories represented. Subpanels were set up to study pistons and piston rings, cylinder liners, cylinder heads and valves, the

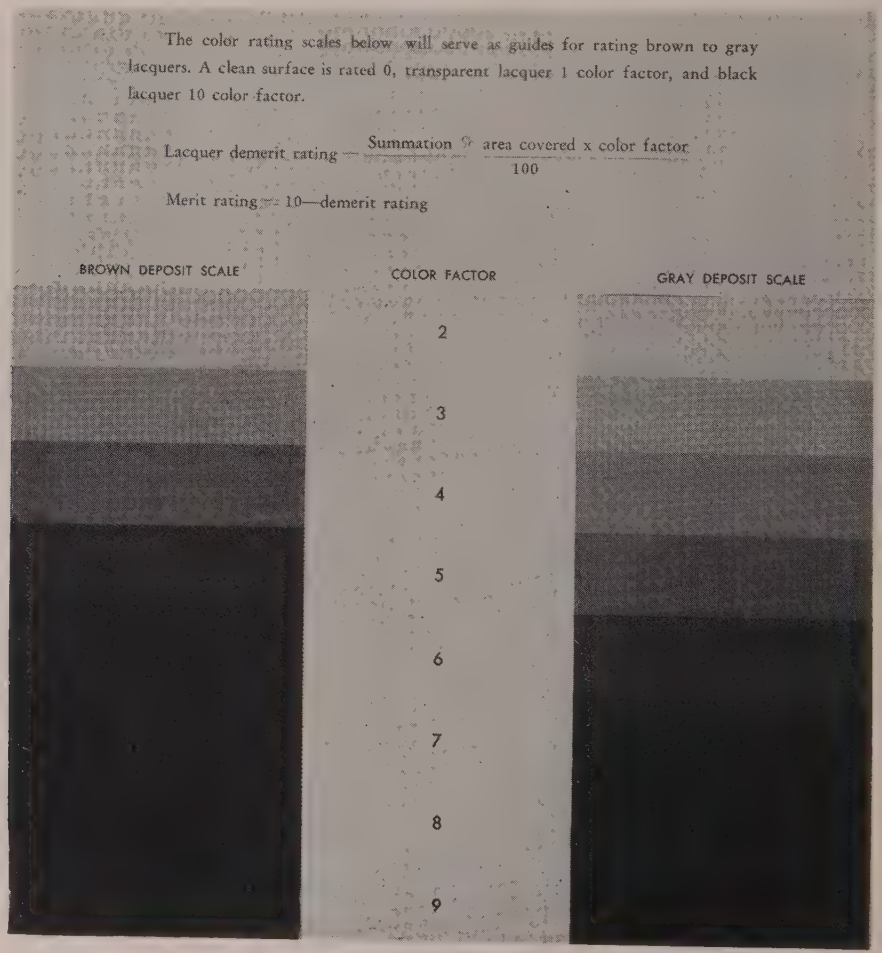

The color rating scales below will serve as guides for rating brown to gray lacquers. A clean surface is rated 0, transparent lacquer 1 color factor, and black lacquer 10 color factor.
Lacquer demerit rating = Summation % area covered x color factor / 100
Merit rating = 10—demerit rating
BROWN DEPOSIT SCALE
COLOR FACTOR
GRAY DEPOSIT SCALE
2
3
4
5
6
7
8
9

Fig. 1 — CRC lacquer rating scale, sample page from "CRC Diesel Engine Rating Manual." Color scales shown are reproductions only and are not to be used for rating. Actual color scales are included in the Manual

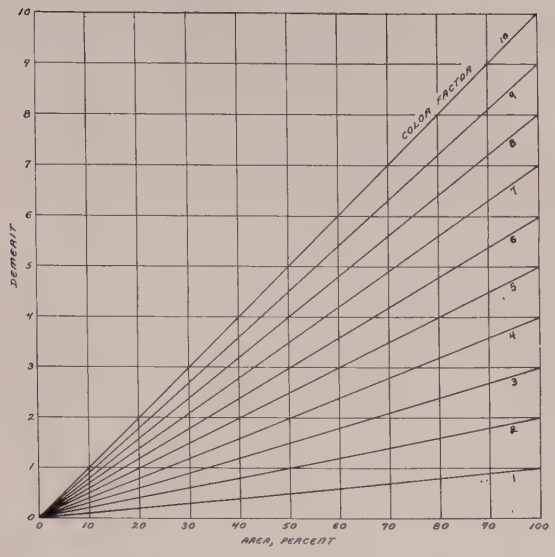


Fig. 2 — Rating scale for lacquer deposits (use CRC lacquer rating scale to obtain color factor)

Fig. 3 — Scratch gage

Fig. 4 — Cleanest possible "A" depth of deposit (deposits less than these illustrated examples should be rated as "clean")

lubrication system and bearings, and the fuel system. It was decided that the objective would be to prepare a manual which could be made available to industry. This manual would not be a system per se, but would be the basis for an engine rater to evaluate any components significant to his test program.

From this time through October, 1957, the subpanels were busy preparing their proposals and reviewing their results with the Rating Systems Panel. Much give-and-take was necessary between the subpanels and the main Panel in order to arrive at a general approach to rating which could be considered compatible. It was decided at this time that ratings for lacquer, sludge, carbon deposits, and ash deposits could be similarly considered regardless of location in the engine.

The subpanel evaluating means of rating pistons made an extensive collection of photographs to illustrate various degrees of piston skirt lacquering. After an exhausive study of these photographs, it was decided diesel-engine pistons could best be evaluated for lacquer deposits by utilizing an area demerit basis and color gradations of brown and gray from clean to black. As illustrated in Fig. 1, a brown scale and a gray scale of 10 graduations were established, 0 being clean and 10 being black. By matching a lacquer deposit condition with either the brown or gray scale, compromising intermediate hues between the two scales to identify shade, a numerical color factor can be established. By utilizing a nomograph (Fig. 2), a demerit value for the surface may be established on a per cent area basis.

The subpanel studying deposits in oil systems considered various means for evaluating thickness and texture. It was decided that the scratch gage developed by the CRC Engine Deposit Rating Panel of the CRC Motor Engine Varnish and Sludge Group was suitable for diesel engines (Fig. 3). A procedure for establishing a volume factor which furnishes a weighted interpretation of the deposit was created. Two thicknesses less than the minimum scratch gage

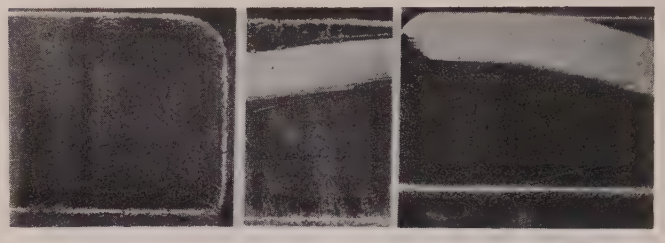


Fig. 5 — Cleanest possible "B" depth of deposit (deposits less than these illustrated examples should be rated as "A" depth)

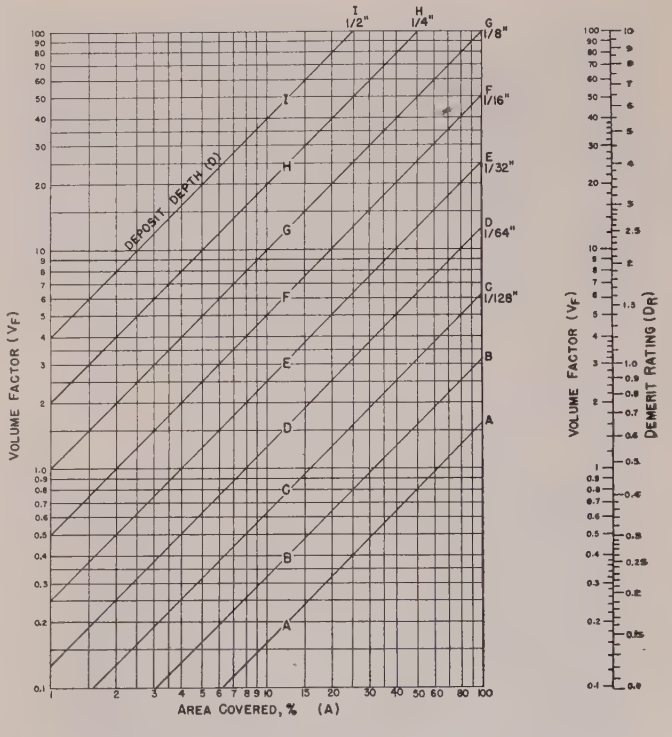


Fig. 6 — Rating scale for sludge, carbon, and ash deposits by CRC volume factor technique (0 $V_F = 0$ D_R = clean over 100% area; 100 $V_F = 10$ D_R = 1/8-in. deposit over 100% area; $V_F = 8$ AD; $V_F = 3$ $D_R^{1.523}$)

thickness (1/128 in.) are recognized and identified as Thickness A and Thickness B. They appear typically as in Figs. 4 and 5. Total surface is rated by averaging the thickness of deposit on a per cent area basis by use of the graph illustrated in Fig. 6. This rating is known as the "CRC Volume Factor," which in turn may be converted to a demerit rating by use of the nomograph.

At the January, 1958, meeting of the Panel, the Editorial subpanel was created to collate the proposals of the subpanels. One of the principal purposes of the Editorial Subpanel was to establish a uniformity of language and rating techniques for all engine parts affected by a similar type of deposit or wear. In June, 1958, the first draft of the Manual was circulated and reviewed in detail at the meeting of the Rating Manual Panel that month in Atlantic City. The product of the Editorial Subpanel was acceptable to the Panel, which decided that the Manual should be tested for general industry acceptability by arranging for trial use in several laboratories.

Arrangements were made to prepare a limited number of draft copies of the Manual for trial rating. The Engine Deposit Rating Subpanel was created to conduct trial engine rating; and in the period of September, 1958, through April, 1959, six different laboratories conducted trial ratings applying the portions of the Manual suitable to their particular interests. Three laboratories which were unable to perform trial rating studied the Manual with respect to their experience and offered comments and recommendations.

At a meeting in June, 1959, at Atlantic City the Engine Deposit Rating Subpanel met, reviewed their experience in rating, and made proposals for improvement in the Manual to the Editorial Subpanel. The Manual was found to be suitable in form and applicable to its purposes. Only minor changes, all editorial in nature, were found necessary, and the Editorial Subpanel completed final revision of the Manual for submission to the CRC-Diesel Committee at that time. At the October, 1959, meeting of the Group on Full-Scale Test Techniques and the CRC-Diesel Committee in Chicago, it was approved and released for distribution.

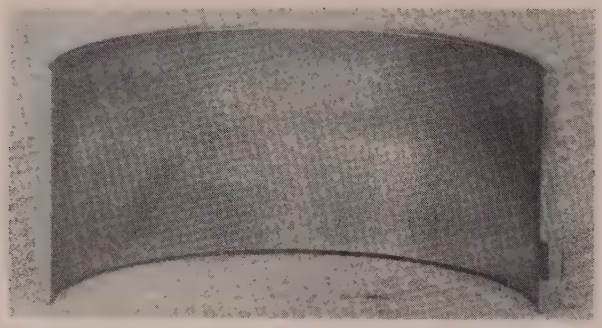


Fig. 7 — Bearing corrosion

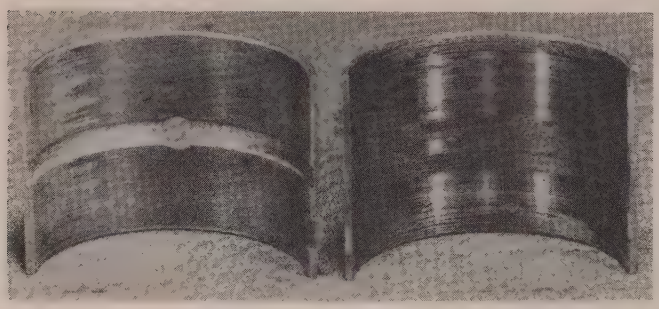


Fig. 8 — Bearing wiping

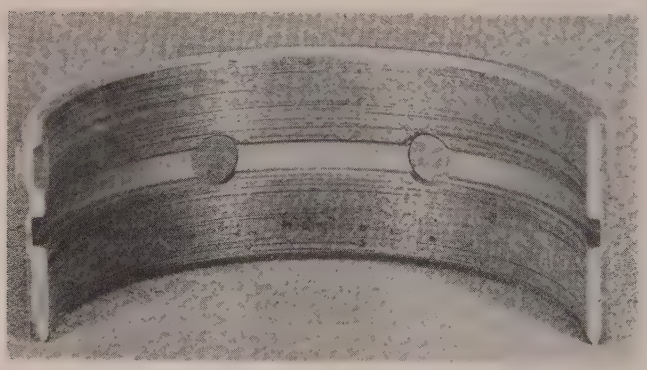


Fig. 9 — Bearing abrasion (dirt in lubricating system)

The Manual

The Manual, as finally evolved, permits any laboratory to define in universally accepted terms those engine conditions significant to performance, life, reliability, and maintenance. It is not anticipated or intended that the Manual will be used in its entirety. Users can take from the Manual techniques for rating portions of the engine within their particular interest. In cases of detailed study of specific engine components, the Manual may only act as a guide toward taking more detailed information than indicated. It is not the purpose of the Manual to evaluate, but rather to define conditions. Significance of conditions observed must be interpreted by the individual laboratory. Reasonable comparability of results between laboratories should be accomplished from use of this Manual.

The Manual is divided into two portions, "General Considerations" and "Detailed Rating Techniques." Terminology and general concepts applicable to engine rating as a whole are kept under "General Considerations." The importance of maintaining a complete record of test conditions is established, along with identification of those conditions which should be recorded. The section on "Detailed Rating Techniques" is devoted to description of rating techniques to be specifically applied to individual parts of detailed areas of a part under consideration. Engine areas included under this section include pistons (including rings), cylinder liners, cylinder heads, valves and valve gear, oil system and airbox deposits, bearings, crankshafts, piston pins, filters, screens, gearing, injection equipment, and engine turbochargers. Under each of these sections, a detailed breakdown is given indicating the applicable technique to be used for rating specific areas of the part mentioned in the particular section. Where advantageous, photographs are included as an aid to the rater. Extensive use of such photographs is made for rating of valves, lifters, gears, and wrist pins, and describing bearing conditions.

The "Album of Reference Photographs — Gear Tooth Condition" prepared by the CRC Gear Lubri-

cants Group, Vehicle Lubricants Division, Coordinating Lubricant and Equipment Research Committee, was found to meet the requirements of identifying gear tooth conditions, and is incorporated in this Manual. Only slight modification to CRC Manual No. 4, "CRC Techniques for Valve Rating" prepared by the Engine Deposit Rating Panel of the Engine Varnish and Sludge Group, was required to make this suitable for rating diesel engines. However, no suitable set of photographs defining bearing deterioration was available, and the Subpanel obtained illustrations of typical conditions from various laboratories. From these, the most suitable were selected for inclusion in the Manual (Figs. 7–11).

The final portion of the Manual includes suggested forms to be used for rating diesel engines for deposits and wear. These forms are typical and must necessarily be modified for test conditions, test objectives, and the make and type of engine involved in a specific test program.

Use of Manual

The fact that a common system of terminology and standardized technique for rating engines is desirable and even necessary is a fact which almost anyone working in this field would be willing to admit. However, for this common language to be of value, it is necessary that laboratories working in this field must use the system so evolved. The question will probably arise as to how the CRC Diesel Engine Rating Manual may fit into a company system which has been well established over a period of years and which apparently has fulfilled the needs of the particular organization that conceived it. While local systems may have fulfilled the need, we believe that it has been apparent to users of such local systems that at times they were operating at a disadvantage when attempting to exchange information or data with workers in other laboratories. When one looks at the CRC Manual, he will in all probability see areas of familiarity; and it is quite likely that he will also see areas of unfamiliarity. While the inclination may be to disregard the areas unfamiliarity, it should be evident that becoming conversant with such areas will in time make the technique set forth as automatic in application as the procedures which a given rater may have used down through the years.

The data generated will then not only be easily obtained but will also form a reservoir of information which can be compared with data generated by other sources. As is stated in the Introduction to the Manual, it should be realized that use of the techniques outlined therein does not prevent a rater from going to greater detail in examining some certain part or areas of a part. Furthermore, if a laboratory is dedicated to the use of numbers for describing all engine conditions, it is certainly possible for it to establish number scales based on the techniques given in the manual for its own purpose. The fact still remains that, if in generating these detailed data, certain standard procedures as set forth in the Manual are originally used, these basic data can still be used at any time for comparison purposes with other laboratories. Since the Manual does not attempt to evaluate, but rather to define conditions, each user is free to set his own limit as to what is satisfactory or unsatisfactory. In this manner, the Manual, while establishing certain standard definitions and concepts, still allows wide latitude to the user for application of the data to his particular needs.

Conclusion

In conclusion, the Coordinating Research Council is pleased to make available to the engineering fraternity this Manual for rating diesel engines for deposits and wear. It is the product of many of the principal laboratories working in this field and, as such, combines the best judgment of hundreds of engineers. To be of value it must be used. The Group on Full-Scale Test Techniques of the Coordinating Research Council will make necessary revisions in the Manual as required to maintain it in keeping with the state of the art.

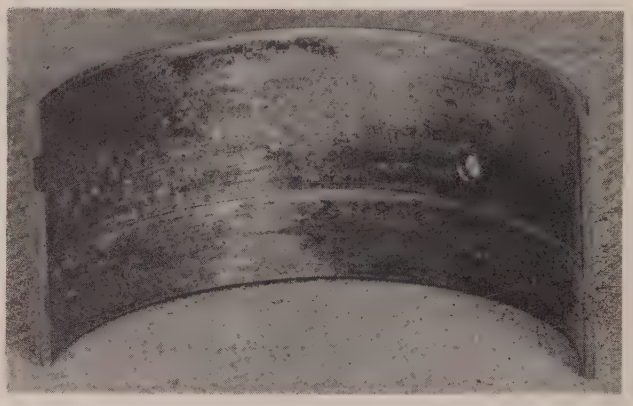


Fig 10 — Embedment of foreign material

Fig. 11A — Fatigue of bearing lining

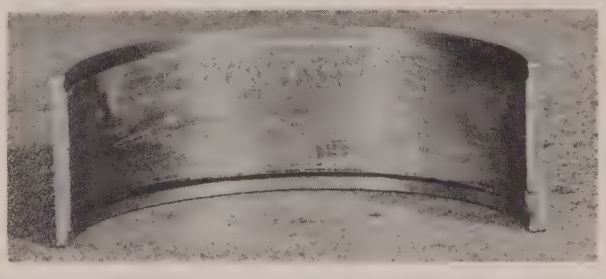


Fig. 11B — Fatigue of bearing lining due to misalignment

Optimum Joint for High-Temperature

JOINT OPTIMIZATION is the compromise of the ideal joint with the most easily producible joint. The design engineer should recognize that his skills are currently at the mercy of the production engineers who must reduce the design to a producible article at a reasonable cost.

This paper discusses two approaches to the design of joints: the pictorial and the analytical. The former emphasizes the optimizing of one type of structure; the latter interprets optimization as achieving the best load transfer at the lightest weight without regard to the particular type of structure.*

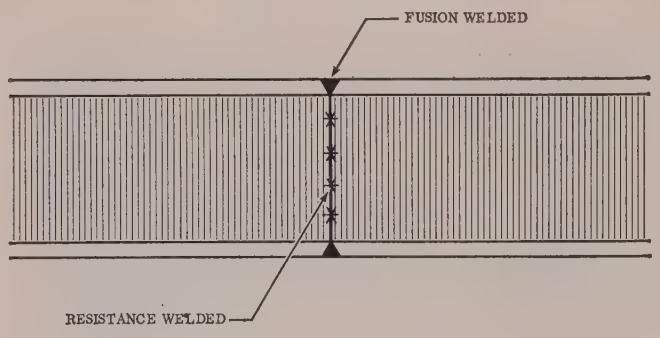


Fig. 1 — Ideal mid-panel joint

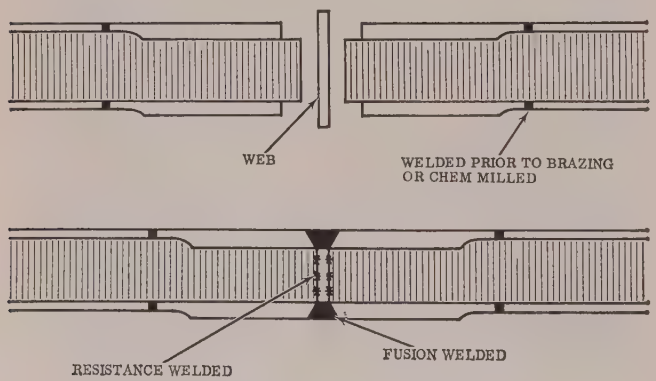


Fig. 2 — Balanced design mid-panel joint

THE EVER-INCREASING demand for high-temperature honeycomb sandwich structures in today's high-speed aircraft and missiles is bringing to light an appalling lack of design information upon which to base such structures. Recent work in optimization of the sandwich itself has established rough guides for application of the brazed cellular structure, but questions inevitable arise as to the manner in which panels are to be joined together without sacrificing efficiency.

There are two basic approaches to the design of joints: the pictorial design approach, used extensively, and the analytic design approach, used less because of lack of time. The pictorial design procedure is concerned mainly with optimizing one type of structure, such as honeycomb sandwich. Analytic design differs in that optimization is primarily interpreted as achieving the best load transfer at the lightest weight, irrespective of the type of structure necessary to accomplish the transfer. This paper briefly discusses both approaches. Regardless of which is taken, one point emerges that deserves emphasis — producibility is the major governing factor influencing the final selection.

Pictorial Design

Recognizing the all-honeycomb joint design as a valid approach to the design problem of joining panels, the discussion will initially center on optimization of this method, to be followed by modification of these designs with other filler materials locally placed to obtain joint optimization.

Mid-Panel Joint — A mid-panel joint must be designed to carry bending, shear, compression, and tension loads. The shear loads require connection of the panel cores, since the core is the only element able to transmit shear forces perpendicular to the panel. For applications not involving shear loads, joining the cores is not mandatory, but is very desirable for joint stabilization. Fig. 1 shows the ideal joint, with the cores and faces welded together, providing necessary ties for proper load transfer. The disadvantage of this joint is that it cannot be produced because of inaccessibility of the core for welding.

A compromise between these extremes should be producible and analytically satisfactory. Fig. 2 represents such a joint. The figure also shows the

* Paper presented at SAE National Aeronautic Meeting, Los Angeles, Oct. 7, 1959.

DESIGN

Honeycomb Panels

Frank J. Filippi and Boris Levenetz
Solar Aircraft Co.

joint before the welding operation as prepared for joining. The now accessible core is welded to the web, the joint is closed, and a fusion weld picks up the faces and web. The addition of the web guarantees a good shear joint for the core, while the thickened facing compensates for the loss in strength from the fusion welding operation. Although some deformation of the core is encountered, good shear transfer is still possible.

Where welding access is from one side only, the joint becomes more cumbersome. A solution is presented in Fig. 3 that is satisfactory except for the incumbent weight penalty. The panels are brazed with the edge members extended beyond the net part size and trimmed to final size. Within limits, growth compensation is possible with this joint. Welding the cover plate to the panel as the last operation completes the joint.

A mechanical joint, admittedly heavier yet, is shown in Fig. 4. Essentially a tongue-and-groove joint, the male portion is a sandwich bar of densified core and faces designed to carry the longitudinal loads of the anels. The bar may be mechanically attached, or brazed in place as shown. The mating portion of the joint is brazed with a U-channel edge member through which the mechanical attachment is made.

Panel-Rib Joint — The panel-rib joint must maintain the integrity of the original panel and transfer any additional loads, in most cases a shear-flow, from the rib into the panel, where this load is distributed between both faces. Fig. 5 shows a joint which theoretically would meet these requirements.

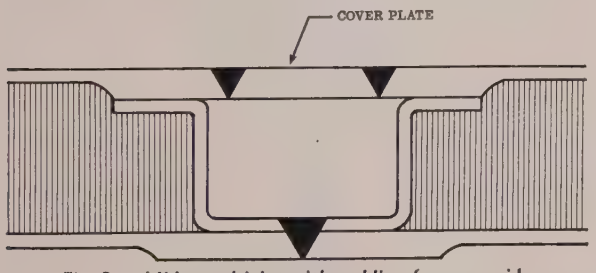


Fig. 3 — Mid-panel joint with welding from one side

The corrugation transfers the rib loads into the outer face of the panel and also makes a link between the honeycomb cores. This joint is plainly difficult to produce because of the accessibility and welding problems.

The design proposed in Fig. 6 meets the design requirements in a producible fashion. This design suggests producing the brazed joint without the web, attaching it by subsequent welding. If this welding can be effected, the panel becomes essentially flat, greatly simplifying the braze fixturing and production sequence. The doubler shown on the drawing may be seamwelded, brazed, or chem-milled as an integral part of the facing. A variation of this

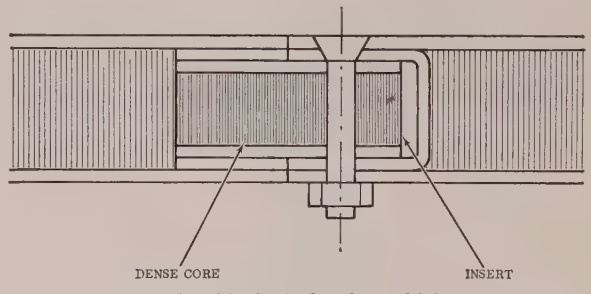


Fig. 4 — Mechanical mid-panel joint

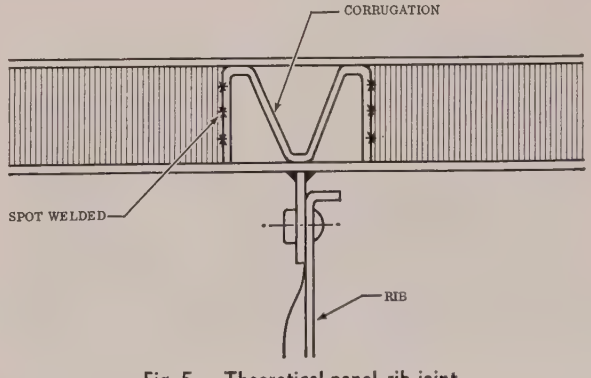


Fig. 5 — Theoretical panel-rib joint

design (Fig. 7) depicts attachment of a honeycomb sandwich rib to the panel. The U-channel is welded to both pieces to effect the joint with a minimum of producibility problems.

A mechanical joint, if acceptable for the application, is quite simple (Fig. 8). The rib is directly attached to a densified core insert which provides reinforcement for the countersink and for the bolt-crushing forces. The doubler between rib flange and panel provides better distribution of compression and shear loads.

Spar-Panel Intersection — The function of this joint is to transmit shear loads from the panels into the spar and to assist the spar in carrying the bending loads by loading an effective width of the panel in tension or compression. Spar caps are shown schematically, and are not necessarily solid. To clarify the attachment problem under discussion, the spar cap detail was not included. The theoretical joint in Fig. 9 incorporates all necessary elements for correct load transfer. Unfortunately, the unreasonable welding requirements make it almost impossible to produce with current knowledge.

A better design is one which incorporates an independently produced spar to which the panels are welded. Fig. 10 is a suggested design for this joint.

A shear web is introduced between the sandwich and the spar as previously discussed before the elements are welded together. This layout shows a honeycomb sandwich spar web riveted to the spar caps with high-shear rivets. If welding from one side only is necessary, a method similar to one previously described is applicable (Fig. 11). The shear transfer is over the short flange of the edge member and cap flange which must have the rigidity to take the shear load in bending. The joint is arranged so the cover plate takes forces only in the plane of the outer facing sheet.

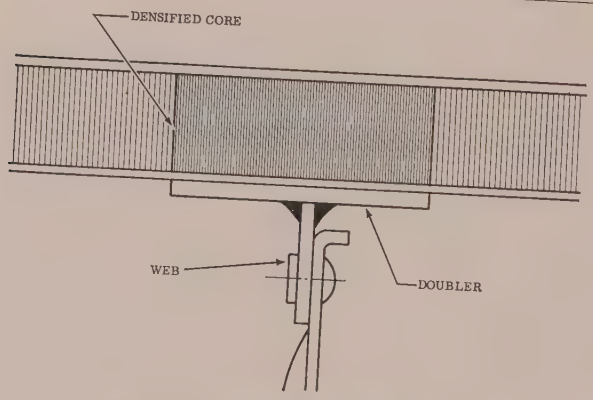


Fig. 6 — Balanced design panel-rib joint

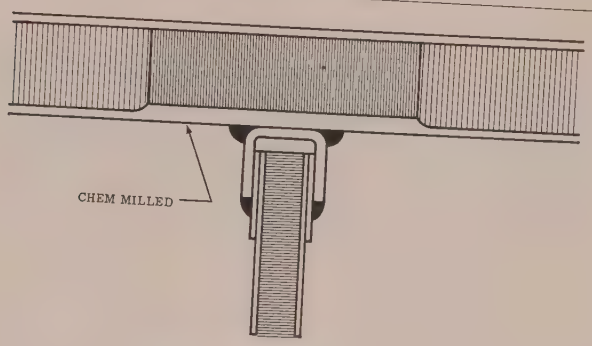


Fig. 7 — Panel-rib joint with honeycomb sandwich rib

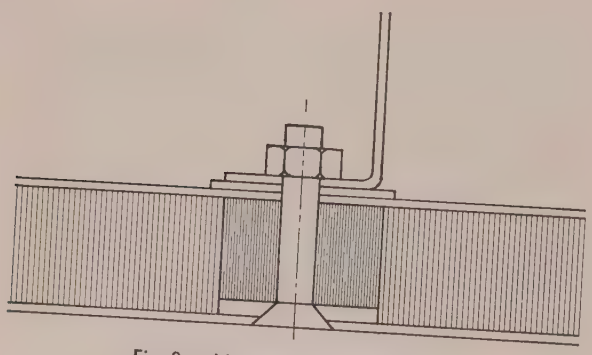


Fig. 8 — Mechanical panel-rib joint

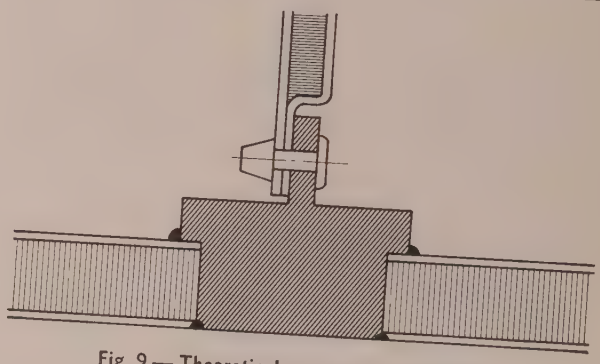


Fig. 9 — Theoretical spar-panel intersection

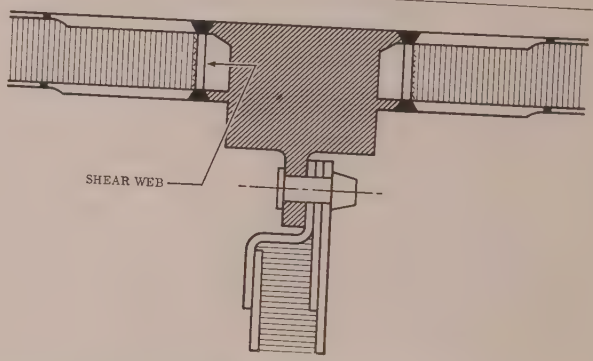


Fig. 10 — Balanced design spar-panel intersection

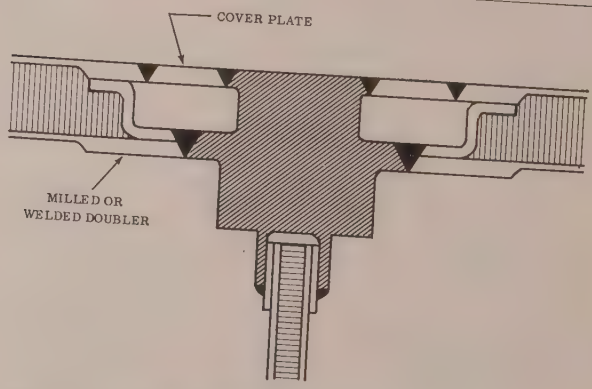


Fig. 11 — Spar-panel intersection with welding from one side

Again reverting to a mechanical joint, Fig. 12 is a composite of two types of panel connection to the spar cap. The double joint takes all kinds of loads effectively, while the single joint is able to take shear perpendicular to the panel and loads in the panel of the outer panel facing only. Another interesting variation of a mechanical spar joint enables a simplified approach to joining the web. Fig. 13 shows a constant-cross-section spar cap incorporated into the basic sandwich. To account for variations in the bending moment, an additional cap strip of varying thickness is placed between the panel and the spar web. The web is then connected to the composite spar cap.

Edge Member Attachment — The edge member configuration is very important because it fixes the types of loads that can be transferred into the panel. Consider three loads, P_x, P_y, P_z acting on a typical Z-section closure (Fig. 14). The load P_x is distributed into the two faces in a manner directly dependent on the stiffness of the Z-section and the core shear modulus in this direction. P_y poses no problem because of the great stiffness of the Z-section in this direction, and it may be assumed that load is transferred equally into both faces. Load P_z is transferred into the vertical web of the Z-section and thus into the core.

Greater control of producibility can be obtained by the design in Fig. 15. Some measure of control over P_x distribution into the two faces is possible with more efficiency than an equivalent-weight Z-section joint permits. Load P_y no longer can be assumed to transfer equally into both facings, but favors the outer facing in proportion to the core shear modulus in that direction. Transfer of P_z now occurs along an undesirable path, placing the core-doubler brazed bond in tension. In addition to the marginality of having a brazed joint in tension, the fatigue characteristics make a joint loaded in this manner exceedingly poor. Obvious, P_z acting in the opposite direction presents a much superior condition.

Fig. 16 is a typical U-channel joint better able to transfer the loads P_x, P_y, and P_z. As a symmetrical

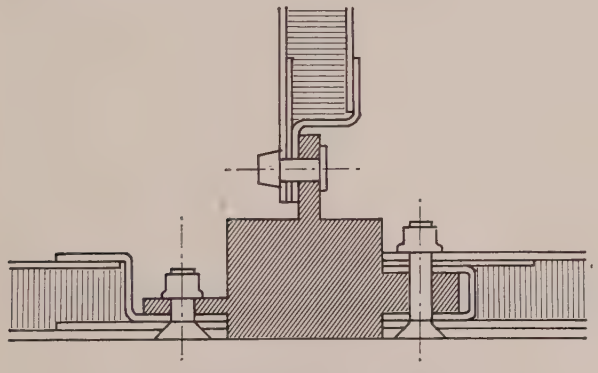


Fig. 12 — Mechanical spar-panel intersection joints

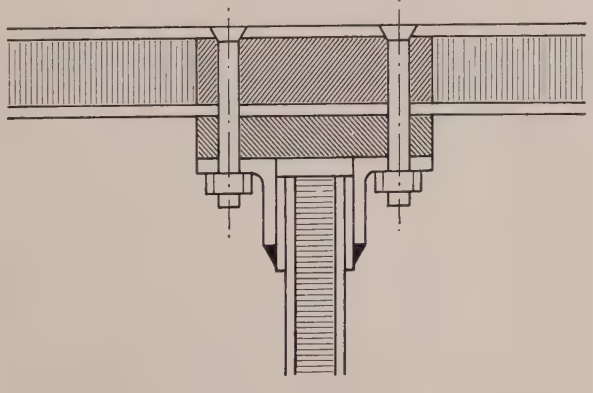


Fig. 13 — Mechanical variation of spar-panel intersection

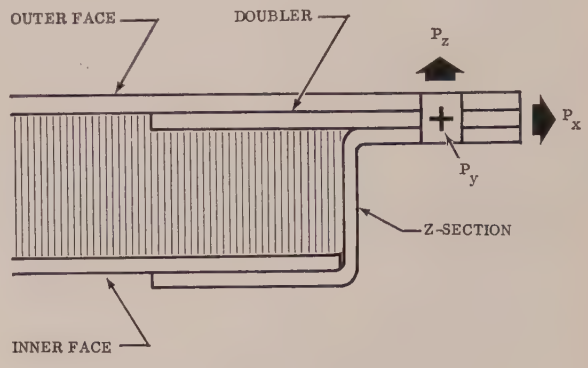


Fig. 14 — Z-section edge member loading

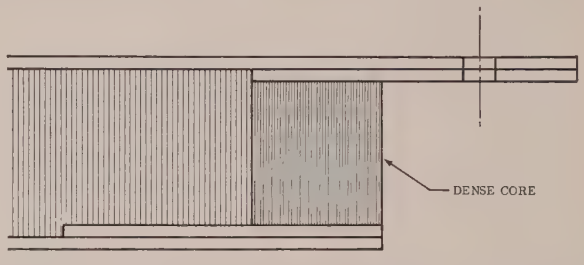


Fig. 15 — Improved edge member producibility, Z-shaped

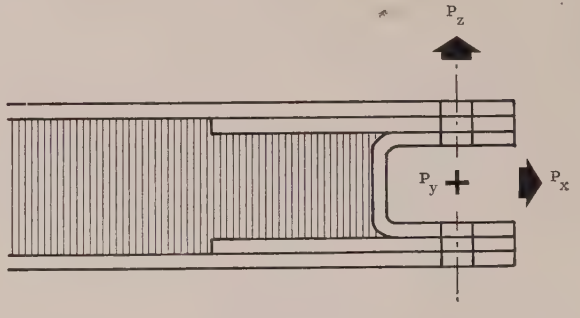


Fig. 16 — U-channel edge member loading

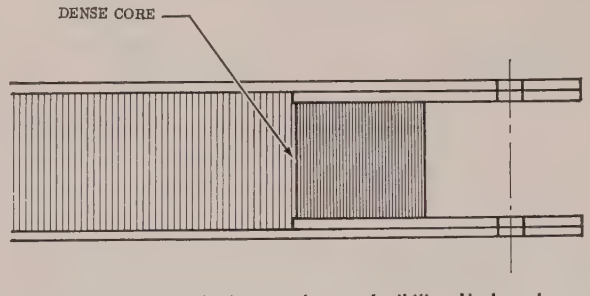


Fig. 17 — Improved edge member producibility, U-shaped

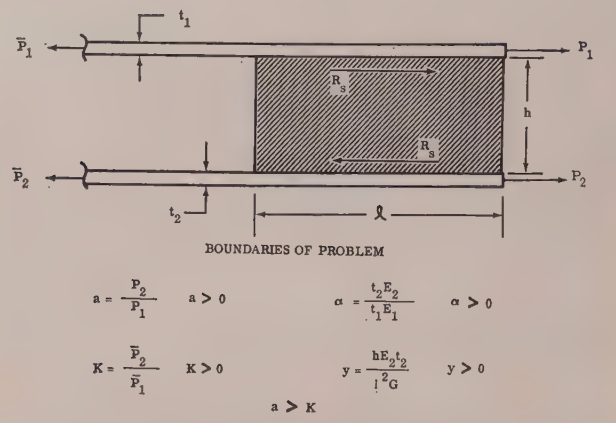


Fig. 18 — Schematic of honeycomb joint

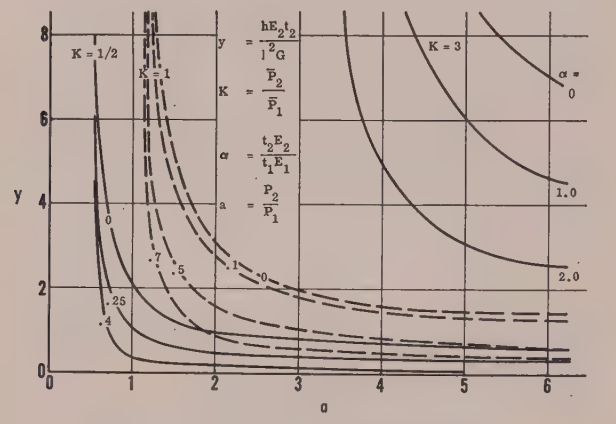


Fig. 19 — Shear modulus required by joint

attachment, the edge member may be assumed to introduce P_x and P_y equally to the facings without regard to channel stiffness. The U-channel provides a shear web for introducing P_z to the core, but otherwise serves little useful purpose. The major problem in production of such a joint — the extremely close tolerance imposed on the formed sheet-metal U-channel — is eliminated by substitution of dense core for the channel in the double area as shown in Fig. 17. The shear transfer of load P_z is much more favorable than with the Z-section because the dense core-doubler bond can always be placed in compression to introduce this load.

Analytic Design

Shanley[1] has shown that joints can account for a large portion of the total structural weight of an aircraft. He notes that the important point brought out by his analyses is that the weight increases associated with joints and nontapered structures are likely to be at least as great as the weight of the optimum structure itself. In his simplified analysis, the weight increase due to joints amounts to 58% of the basic structure and the increase due to nontapered sections amounts to 29%. This, then, justifies vigorous treatment of the joint problem.

Optimization of joints can be accomplished by standard procedures, with the joint treated as a structure with certain imposed loads. One approach will be illustrated which, although not intended to

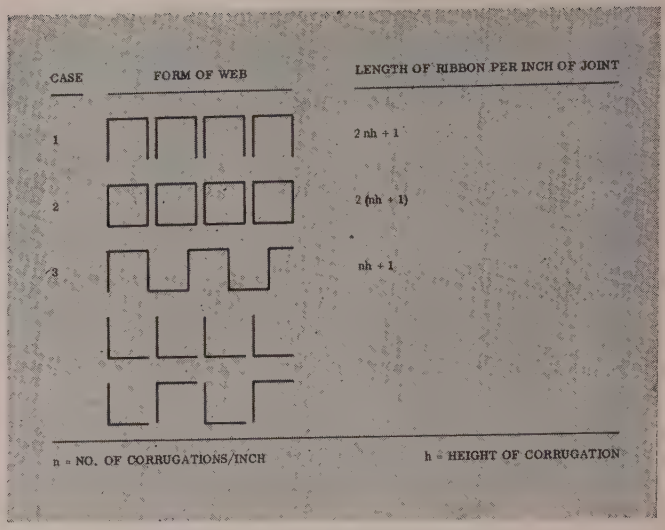


Fig. 20 — Straight shear webs considered

be applicable to all honeycomb panel joints, illustrates several general techniques and demonstrates the amount of material and weight saving that can be had by careful design.

The joint chosen is a simple but important one — a joint between sandwich panel and supporting structure, where the facings of the sandwich are loaded in tension. Let us assume, however, that the loads in each face are not equal in the panel or in the joint. This is the usual case in the joint, because the reacting structure has varying rigidity and will usually resist loads in one sandwich facing more effectively than those in the other facing. The load picture for such a joint is as shown in Fig. 18 along with the imposed boundary conditions. The equations of statics and compatibility of deformations yield the following set of equations:

$$\delta_1 = \frac{P_1 l}{t_1 E_1} + \frac{R_S l}{t_1 E_1}$$

$$\delta_2 = \frac{P_2 l}{t_2 E_2} - \frac{R_S l}{t_2 E_2}$$

$$\delta_2 - \delta_1 = \frac{h R_S}{lG}$$

where:

δ = Deflection, in.
P = Load in facing in joint, lb/in.
l = Width of joint, in.
t = Thickness of facing, in.
E = Modulus of elasticity, psi
R_S = Shear load, lb/in.
h = Core height, in.
G = Modulus of rigidity, psi
Subscripts refer to the panels

Solving for R_S:

$$R_S = P_1 \frac{a - \alpha}{y + \alpha + 1} \qquad (1)$$

where:

$$a = P_2/P_1$$
$$\alpha = t_2 E_2 / t_1 E_1$$
$$y = hE_2 t_2 / l^2 G$$

The loads P_1 and P_2 can be related in many ways,

[1] "Weight-Strength Analysis of Aircraft Structures," by F. R. Shanley. Pub. by McGraw-Hill Book Co., Inc., New York, 1952.

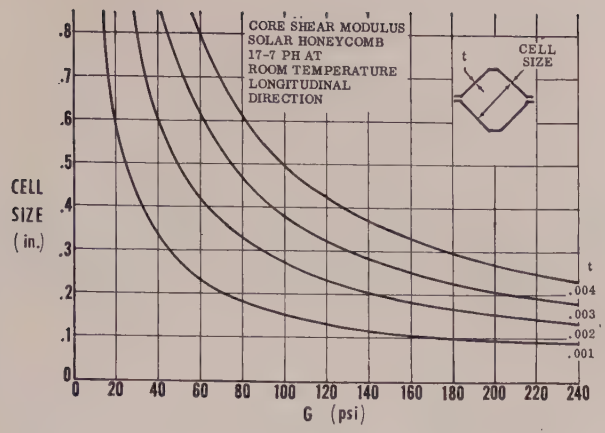


Fig. 21 — Core shear modulus for honeycomb

depending on loading conditions, temperatures[1], and the like. Defining $\bar{P}_2/\bar{P}_1 = K$, where $\bar{P}$ is the load in the panel facing in lb/in., and noting that:

$$\bar{P}_2 = P_2 - R_S$$

$$\bar{P}_1 = P_1 + R_S$$

yields:

$$K = \frac{P_2 - R_S}{P_1 + R_S}$$

and:

$$R_S = \frac{P_2 - KP_1}{K+1} = P_1 \frac{a - K}{K+1} \qquad (2)$$

Combining Eqs. 1 and 2 yields:

$$R_S = P_1 \frac{a-K}{K+1} = P_1 \frac{a-\alpha}{y+\alpha+1}$$

and:

$$y = \frac{(a+1)\ (K-\alpha)}{a-K} \qquad (3)$$

Eq. 3 defines the value of $y = \frac{hE_2t_2}{l^2G}$ required to maintain equilibrium in a joint with any ratio of applied loads K, and of reacting loads, a. Eq. 3 is plotted in Fig. 19 for several values of K. Limiting conditions can be noted as follows:

1. When $a = K$; *that is, when* $P_2/P_1 = P_2/P_1$, $y = \infty$ and the required shear modulus = 0. This means that no shear load is required in the joint, and $R_S = 0$.

2. When $K = \alpha$, $y = 0$ for all values of a, so the G required is infinite, or the joint width is infinite. This condition implies different face thicknesses in the joint area, that is, a doubler. This is true only when $K = \alpha$ and $K \neq a$. When $K = a$, the first condition applies instead.

The relationship of Eq. 3 defines the core shear modulus required to transfer a given load from one face of a sandwich to the other — that is, to change $\bar{P}_2/\bar{P}_1$ *to* P_2/P_1 across a given joint width l. The core shear modulus available in two conventional core materials will now be determined. The two configurations to be considered are honeycomb core and straight shear webs. The shear webs can occur in several forms, as shown in Fig. 20. Since the length of ribbon per lineal inch of joint varies for each of these forms, the density of each form also varies.

Honeycomb Core

Examine first the honeycomb core. Relationships are needed for ultimate shear strength, core shear modulus, and core density.

Ultimate Shear Strength — A formula for calculating core shear strength is:[2]

$$F_S = K_2 \frac{g_a^2}{rg - g_a}$$

where:

g_a = Density of honeycomb core, lb/cu in.
r = Periphery of a honeycomb cell divided by twice the largest diameter. For the square cell, $r \approx 2$

If K_2 is treated as a constant for all cores, it can be evaluated by test of one core. Test data shows the ultimate shear stress associated with 3–15 core, 17–7 PH material at room temperature to be approximately 850 psi. Using $g_a = 0.0046$ lb/cu in., $g = 0.286$ lb/cu in., and $r = 2$, yields $K_2 = 22.85 \times 10^6$ psi.

The relationship for core shear strength can then be written:

$$P_S = 22.85 \times 10^6 \frac{lg_a^2}{rg - g_a} \qquad (4)$$

where:

g = density of material from which core is made, lb/cu in.
P_S = allowable shear load in joint, lb/in.

Core Shear Modulus — Haener[3] summarizes an analysis performed at our company to derive an analytic expression for the core shear modulus of any regular, cellular, honeycomb core in any direction at any temperature. This equation retaining the notation of Haener is:

$$\frac{G_F\rho_S}{G_S\rho} = \frac{(R + \cos\theta)^2 \sin^2\theta}{(1+R)[(1+R)\sin^2\theta + \{(R+\cos\theta)^2 - (1+R)\sin^2\theta\}\sin^2\phi]}$$

where:

G_F = Core shear modulus
ρ_S = Density of core material
G_S = Shear modulus of core material
ρ = Core density
$R = \frac{q}{S}$
q = Length of wall of double thickness in core
S = Length of wall of single thickness in core
θ = Angle between ribbon direction and cell side S
ϕ = Angle between applied force and ribbon direction

For Solar honeycomb core, in the longitudinal direction, $\theta = 45$ deg, $\phi = 0$ deg, $q = 0.04$ in., and

$$G_F = G_S t_f \frac{A + 0.02828}{(A - 0.02828)\ (A + 0.02828)} \qquad (5)$$

where:

A = Cell size of honeycomb core

A curve of this equation is shown in Fig. 21 for 17–7PH, TH 1050 at room temperature.

Core Density — Kaechele[4] presents an approxi-

[2] "Sandwich Construction for Aircraft." U. S. Department of Defense Publication ANC-23, Vol. 2, 1955.

[3] "Calculation of Core Shear Modulus of Honeycomb Sandwich," by J. A. Haener. Report, Solar Aircraft Co.

[4] "Minimum-Weight Design of Sandwich Panels," by L. E. Kaechele. Report RM-1895, The Rand Corp., 1957.

mate formula for core density:

$$g_a = \frac{8}{3}\frac{t_f}{A} g \qquad (6)$$

where:

t_f = Thickness of the corrugation foil

Shear Web

Eqs. 4, 5, and 6 define the weight of honeycomb core necessary to produce a given core shear modulus, and also the maximum shear load to which a honeycomb core can be used. This same procedure may be used for straight shear webs.

Ultimate Shear Strength — The shear webs will fail by buckling, which is defined by the classical formula:

$$\tau_{cr} = K \frac{\pi^2 E t^2}{12(1-\nu^2)h^2} \qquad K = 9.4 \text{ when } \frac{l}{h} = 1$$

$$6.6 \text{ when } \frac{l}{h} = 2$$

$$5.35 \text{ when } \frac{l}{h} = \infty$$

where:

τ_{cr} = critical shear stress, psi

ν = Poisson's ratio

The total shear load carried by the webs per inch of joint is then

$$P_{cr} = nK \frac{\pi^2 E t^2 l}{12(1-\nu^2)h^2} \qquad (7)$$

where:

n = Number of straight shear webs per inch of joint

Shear Modulus — The shear modulus for the straight webs is that of the web material. For 17-7PH-TH1050 steel at room temperature, $G = 11.3 \times 10^6$ psi.

Weight of Shear Webs — The core weight per inch of joint is the length of ribbon per inch of joint, multiplied by $lt_f g$. The ribbon length per inch of joint is shown in Fig. 20 for each type of core considered. These relationships define, for the straight shear web, the core weight required to carry a given load in the joint and the maximum core shear modulus available.

Fig. 22 presents curves showing the relative efficiency of honeycomb core and shear webs in carrying joint loads. The curves are calculated for $h = l$ and 17-7PH/TH1050 material at room temperature.

Eq. 7 simplifies to

$$t_f = 6.26 \times 10^{-5} \sqrt{\frac{P}{n}}$$

and Eq. 4 simplifies to:

$$g_a = 1.58 \times 10^{-4} \sqrt{P} - 2.19 \times 10^{-8} P$$

Fig. 23 demonstrates clearly that large variations in core weight can occur at any given load. It shows further that, in theory, honeycomb core is not as efficient as straight shear webs in joints for any value of joint load. It must be noted that the shear web curves will be cut off at the lower loads by minimum-gage considerations.

In these cases it might be advisable to consider different materials or configurations for the joint core. These same equations can be used to determine doubler thicknesses required in the joint area.

Optimization

Both pictorial and analytic design approaches have been illustrated to show the usefulness of each procedure. Pictorial design obviously is a useful tool for comparative screening through instinct and judgment. Unfortunately, subsequent analytic work is often limited to determination of dimensional adequacy of a selected sketch. Such work amounts to little more than design confirmation, although the procedure is often mistakenly called optimization.

The Analytic approach attempts to synthesize a joint from a series of efficient configurations known to have certain load-carrying abilities. The analyst considers the loads in terms of hardware details which he then attempts to combine into an integral, producible structure. To accomplish this goal, it is necessary to calculate all the applicable properties of each detail singly and in combination and compare them. Where several potential designs fulfill design requirements, a weight comparison separates the most effective configuration. The analytic procedure's main weakness is that producibility is relegated to a position of secondary importance.

It is logical to conclude that producibility must be the medium of expression through which pictorial and analytical design can be combined to enable production of economical, efficient parts.

Acknowledgment

The authors gratefully acknowledge the analytic work performed at Solar by M. R. Licciardello, senior engineer, Structures, in support of this paper.

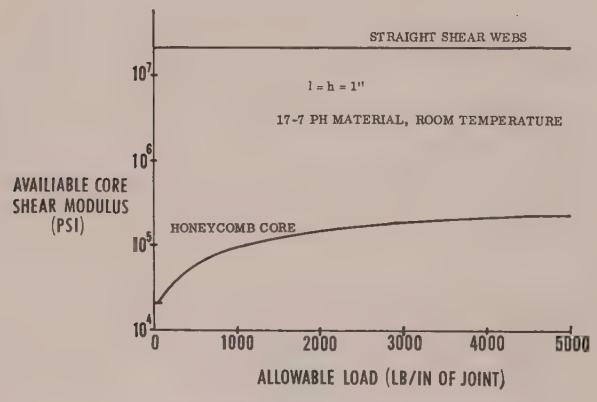


Fig. 22 — Maximum core shear modulus available

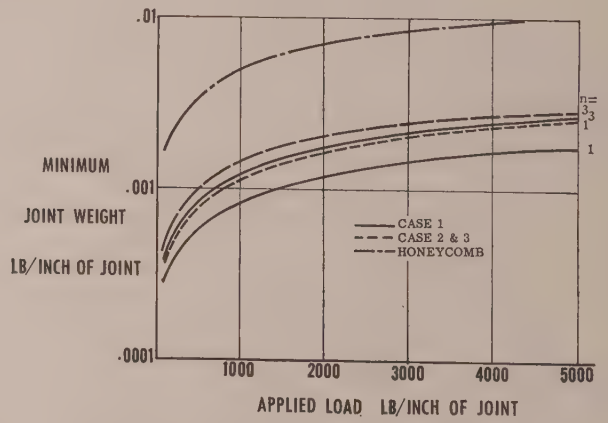


Fig. 23 — Relative weights of panel joints

GMR Stirling Thermal Engine

part of the Stirling engine story

—1960 chapter

Gregory Flynn, Jr., Worth H. Percival, and F. Earl Heffner
Research Laboratories, General Motors Corp.

THE PRINCIPLES of the Stirling engine are not new — they have been known for more than a century. General Motors Research Laboratories' study of the Stirling engine is not new. It has continued for more than a decade. The modern Stirling thermal engine, however, bears little resemblance to its historical predecessors except in principle; and many man-hours of analysis and testing have served to instill the modern engine with subtle qualities that can only be the products of present-day scientific and engineering skill. Most of the credit for this successful application of modern skills must go to the physicists and engineers of the Philips Research Laboratories of Eindhoven, Netherlands.

In November, 1958, N. V. Philips' Gloeilampenfabrieken of Eindhoven and the General Motors Corp. agreed to institute a cooperative program to develop further Stirling engines for American commercial and military applications. Under this program, the first two modern, full-size Stirling engines ever to operate in the United States were tested at the Research Laboratories in September, 1959. These first two engines were built by Philips according to the principles and practices evolved during their 20-year research program, and it is primarily on the basis of the performance of these two engines that General Motors Research decided to make a public introduction of this new prime mover. The Stirling thermal engine described in this paper is a later engine configuration and represents a cooperative design of both Philips and General Motors Research.

The Stirling thermal engine differs from other modern engines in one major aspect — it is an external combustion engine. The fuel, the combustion air, and the products of combustion never enter the engine cylinder to become part of the working gas. On the other hand, the working medium is completely sealed within the active spaces of the engine and passes through the thermodynamic cycle repeatedly to cause the engine to operate. Heat is transferred from the combustion products to the working fluid through the metal walls of a heat exchanger, and heat is rejected from the working fluid through the metal walls of a second heat exchanger.

The Stirling thermal engine will operate from any heat source of sufficiently high temperature. Its efficiency is very high. Its sound and vibration levels are very low. Its power to swept-volume ratio is high, and its power-to-weight ratio is competitive.

Stirling Engine Cycle

The Stirling engine is like all modern engines in that its operation is guided by the processes of an ideal thermodynamic cycle through which its working fluid is intended to pass. Also like most modern engines, the actual cycle of operation is considerably different from the ideal upon which it is based; but in introducing this new engine it is useful to consider the characteristics of the basic Stirling

THIS PAPER discusses the Stirling thermal engine from four points of view:

1. The ideal, thermodynamic point of view, showing the inherent potentialities of the ideal Stirling cycle in comparison to the basic cycles of other engines.

2. The physical engine and its method of operation with respect to the ideal cycle and the limitations of practical mechanics.

3. Performance data from the first modern Stirling engines ever operated in the United States, evaluating the relationship between the new engine and other more familiar engines of similar sizes. This comparative discussion serves to demonstrate the advantages and disadvantages of the Stirling engine and to indicate its proper place in the 1960 family of prime movers.

4. A look backward into the century of history behind the modern engine pointing out significant milestones in the engine's development. Particular attention is paid to Stirling's work, to the engine's American history, and to the significant, recent developments that were instrumental in producing the new Stirling engine*

cycle because the most powerful properties of the basic cycle manifest themselves in the operation of the actual engine. A similar cycle, the Ericsson cycle, will also be considered briefly since many of the historical predecessors of the modern Stirling thermal engine were intended to operate according to that cycle.

The Stirling and Ericsson cycles are shown schematically on the pressure-volume diagrams in Fig. 1. The Stirling cycle consists of isothermal expansions and compressions between constant volume heat additions and heat rejections. The Ericsson cycle substitutes constant pressure heat addition and rejection processes for the constant volume processes of the Stirling cycle.

The first process of the Stirling cycle is an isothermal compression from point 1 to 2. Next, heat is added during a constant volume process from 2 to 3. The isothermal expansion or power stroke then takes place from 3 to 4, and the cycle is completed by the constant volume cooling process from 4 to 1.

Heat is added to the working fluid during the constant volume process, 2–3, and during the isothermal expansion, 3–4. Heat is rejected during the constant volume cooling process, 4–1, and also during the isothermal compression, 1–2.

Modern Stirling thermal engines always have regenerators, for otherwise their thermal efficiency would be very poor. The purpose of the regenerator is to store the heat removed from the working fluid during the constant volume cooling process, 4–1, and return it to the working fluid during the constant volume heating process, 2–3. It can be shown that in the ideal case, equal quantities of heat are removed from and returned to the fluid during these two constant volume processes. In the ideal regenerative Stirling cycle, then, heat from an external source must be added only during the isothermal expansion, and heat must be rejected to an external sink only during the isothermal compression.

The Ericsson cycle is similar, starting with an isothermal compression from 1 to 2, then a constant pressure heat addition from 2 to 3, an isothermal expansion from 3 to 4, and a constant pressure cooling from 4 to 1. Again heat must be added to the working fluid during the constant pressure process, 2–3, and during the expansion, 3–4; while heat must be removed from the working fluid during the constant pressure cooling process, 4–1, and during the isothermal compression, 1–2. Also, a regenerative Ericsson cycle is much more efficient than the nonregenerative cycle; the purpose of the regenerator being to store the heat from the constant pressure process, 4–1, and to use it to heat the fluid during the constant pressure process, 2–3. As in the Stirling cycle, it can be shown that these two quantities of heat are equal in the ideal cycle, so that external heat need be added to the engine only during the isothermal expansion, 3–4, and rejected from the engine during the isothermal compression, 1–2.

Heat is added to both the regenerative Stirling and the regenerative Ericsson cycles at a constant temperature equal to the highest temperature of the working fluid, and it is rejected only at a constant temperature equal to the lowest temperature of the working fluid. For this reason the ideal regenerative Stirling and Ericsson cycles have the same thermal efficiency as the Carnot cycle operating between the same temperature limits. The Stirling cycle is the only basic modern engine cycle that has this inherent capability of producing the ultimate maximum efficiency of the Carnot heat engine.

This brief, qualitative description of the Stirling and Ericsson cycles can be supplemented by consideration of the relationship of these two cycles to the familiar Otto and Brayton cycles and to the ideal Carnot cycle. To facilitate this comparison among the five cycles, temperature-entropy and pressure-volume diagrams are shown in Fig. 2. Some significant numerical data pertaining to these diagrams are given in Table 1.

These diagrams and data were prepared for so-called "Air Standard" cycles for which it is postulated that the working fluid is air with constant specific heats, and that all of the processes are ideal and reversible. Under these conditions, the area

* Paper presented at SAE Annual Meeting, Detroit, Jan. 11, 1960.

within the cycle on the temperature-entropy (TS) diagram equals the heat equivalent of the net work performed per pound of working fluid, and the area within the cycle on the pressure-volume (PV) diagram equals the work performed by the actual mass of fluid for the pressures and volumes plotted. The TS diagram is, therefore, useful for indicating efficiencies, temperatures, and heat flow effects; while the PV diagram serves to allow comparison of the output, pressure, and volume effects of similar sized engines. The TS and PV charts of Fig. 2 are drawn to scale within themselves, but they are not drawn to the same scale with respect to each other. The TS diagram expresses work in heat units; that is, Btu, while the PV diagram shows work in inch-pounds. Drawn to the same scale, the area of the PV diagram would be nearly 10,000 times that of the TS diagram for an equivalent amount of work.

In addition to the general conditions stated above, the following particular conditions were established for the comparison cycles. First, the overall efficiency of all five cycles was made equal to that of a 9.5/1 compression ratio Otto cycle, 59.5%. The cycles all have the same minimum temperature, 540 R; the same minimum pressure, 14.7 psia; and the same maximum cylinder volume, 47 cu in. Furthermore, the Brayton, Stirling, and Ericsson cycles have 100% effective regenerators. The Otto cycle engine does not commonly employ a regenerator, and the Carnot cycle does not need one.

The compression ratio of the Otto cycle was arbitrarily set at 9.5/1 to approximate common practices, and the amount of heat added to the cycle during the process, 2–3, was selected to limit the maximum cycle temperature to the value shown. The efficiency of the ideal Otto cycle depends only upon the pressure ratio, but the output per pound of working fluid varies with the amount of heat added.

The maximum temperature of the regenerative Brayton cycle was arbitrarily set at the value shown, and the pressure ratio selected to give the desired efficiency common to all five cycles. The efficiency of this Brayton cycle operating between given temperature limits approaches Carnot efficiency as the pressure ratio decreases, but the net output per pound of working fluid decreases with decreasing pressure ratio.

The volume ratio of the Stirling cycle was arbitrarily set at two since this approximates a practical value. This volume ratio and the other particular conditions stated above completely defined the Stirling cycle, and the Carnot and Ericsson cycles were designed for the same amount of heat addition and net work per pound of fluid as the Stirling cycle.

These ideal cycles would give engine cylinder volumes of reasonable size and would operate within limits of practical temperatures and pressures, and

Table 1 — Ideal Air Cycle Data

	Otto	Brayton	Carnot	Stirling	Ericsson
Temperature-Entropy Data					
Efficiency, %	59.5	59.5	59.5	59.5	59.5
Minimum Cycle Temperature, R	540	540	540	540	540
Minimum Cycle Temperature, R	3418	2100	1330	1330	1330
Carnot Efficiency, %	84.1	74.3	59.5	59.5	59.5
Net Work, Btu/lb	211	110	37.7	37.7	37.7
Pressure-Volume Data					
Max. Cylinder Volume, cu in.	47	47	47	47	47
Mass of Working Fluid, lb	0.0020	0.0008	0.0020	0.0020	0.0008
Net Work per Cycle, Btu	0.422	0.088	0.075	0.075	0.030
Hp at 3000 cpm, hp	29.9	6.2	5.3	5.3	2.1
Minimum Cylinder volume, cu in.	4.94	6.10	2.48	23.5	9.6
Volume Ratio	9.5	7.8	19.1	2.0	4.94
Piston Displacement, cu in.	42.06	40.90	44.52	23.50	37.40
Hp/cu in.	0.71	0.15	0.12	0.23	0.06
Minimum Cylinder Press., psia	14.7	14.7	14.7	14.7	14.7
Maximum Cylinder Press., psia	884.0	72.4	688.0	72.4	29.4
Pressure Ratio	60.0	4.94	46.7	4.94	2.00

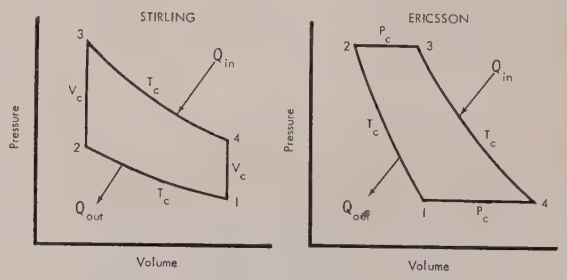


Fig. 1 — Stirling and Ericsson cycles

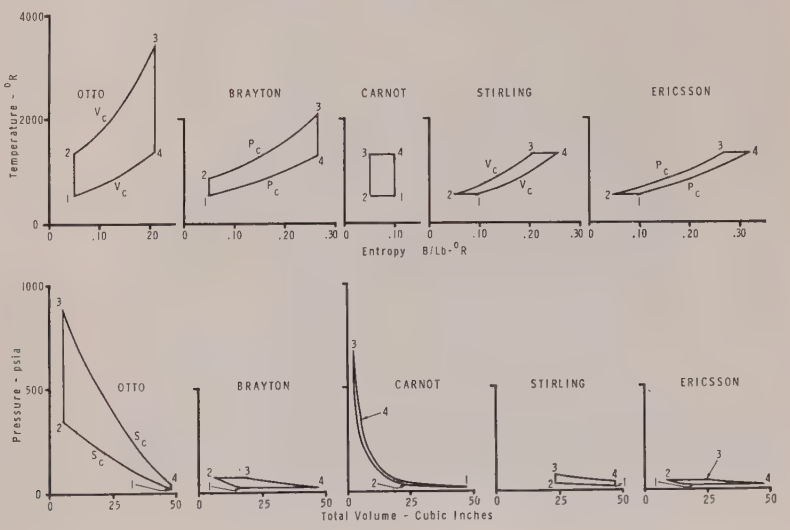


Fig. 2 — Temperature-entropy and pressure-volume diagrams

therefore will serve as a suitable basis for comparison of the inherent basic characteristics of the five kinds of engines.

Referring to the TS diagram and the tabular data pertaining to it, it may be seen that the Otto cycle requires the highest maximum temperature to attain the common efficiency, that the Brayton cycle maximum temperature is next highest, and that the Stirling, Ericsson, and Carnot maximum temperatures are equal and lower. Since the Carnot efficiency is the highest possible efficiency of a heat engine operating between two temperature limits, the Stirling and Ericsson cycles enjoy a significant advantage over the other two with regard to efficiency. This, then, is the first characteristic of the ideal Stirling engine — it operates at the maximum possible efficiency of any engine. A Stirling cycle between the temperature limits shown for the Brayton cycle would have an efficiency of 74%; and between the temperature limits of the Otto cycle, the Stirling cycle would have an efficiency of 84%.

It is reasonable to consider that a gas turbine, or Brayton engine, and a Stirling engine could have the same maximum temperatures since some engine component must operate near this maximum temperature in both instances. The Otto engine, however, has an advantage in this respect in that the maximum fluid temperature occurs only periodically during the cycle and the hottest part of the engine may operate at a substantially lower temperature.

In addition to these temperature and efficiency effects, the TS diagram shows that the Otto cycle produces the greatest amount of work per pound of fluid with the Brayton cycle producing about one-half as much and the other three cycles about one-sixth as much. It should be noted that these output figures are for one pound of working fluid passing through the cycle. The output for engines of the same size are compared on the basis of the PV diagrams.

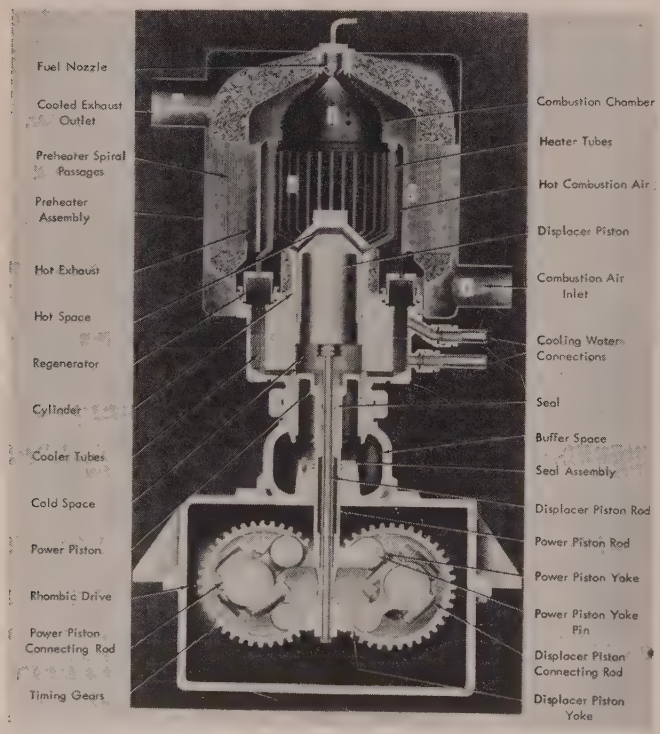


Fig. 3 — Stirling thermal engine schematic drawing

These PV diagrams and the accompanying portion of the data table show the cycle characteristics of five engines each of which has a total maximum cylinder volume of 47 cu in. First, the table shows that the mass of working fluid handled per cycle in the Brayton and Ericsson engines is considerably below that of the other three. This is because the maximum cylinder volume of these cycles must occur at a time when the fluid temperature is higher than the minimum cycle temperature, thereby causing the fluid to be less dense.

This reduced mass per cycle is reflected in the output per cycle and the engine output at a constant speed of 3000 cpm. The Otto cycle engine still enjoys an inherent advantage in specific output with the Brayton, Carnot, Stirling, and Ericsson engines following appreciably behind in the order given. This output advantage of the Otto engine persists even if one considers a four-stroke cycle and allows for the fact that 3000 rpm of the engine would give only 1500 cpm.

The pressures, volumes, and their respective ratios are given in Table 1; and it is seen that both the Otto and Carnot engines attain high cycle pressures while the Brayton, Stirling, and Ericsson engines are basically low-pressure ratio engines.

If the Stirling engine were operated at the same maximum cyclic pressure as the Otto engine, its output at 3000 cpm would be six times that shown or 31.8 hp. This is double that of the four-stroke-cycle Otto engine. A similar computation might be made for the Brayton and Ericsson engines; but the results would not be particularly instructive, since highly supercharged Brayton engines are not built, and since the Ericsson cycle is only of historical interest.

The above observations are about all that need be made on the basis of ideal cycle diagrams. Actually, the Brayton engine achieves higher output for its size by employing turbine machinery capable of handling large quantities of fluid per unit time and thereby overcomes the inherent disadvantage of low output per unit fluid quantity. The Ericsson cycle is not used any more, and there is no practical way of making a Carnot engine.

On the basis of this brief discussion of ideal air cycles, the following comparisons between Stirling engines and other common engines may be made:

1. Between given temperature limits the Stirling engine cycle is the only practical engine cycle to attain the thermodynamic maximum efficiency.

2. At the same minimum cycle pressure the specific output of a Stirling engine is 85% of that of the same displacement Brayton engine, and one-third of that of the same displacement four-stroke Otto engine.

3. At the same maximum cycle pressure the specific output of a Stirling engine is more than double that of the same size four-stroke Otto engine.

Inherently, the Stirling engine has high efficiency, but it must handle large quantities of working fluid to attain competitive specific outputs. Since the Stirling thermal engine is a closed-cycle engine which uses one charge of working fluid repeatedly, it can be made to handle a large mass of this working fluid simply by maintaining a high mean pressure in the working spaces. Thus, in effect, the

Stirling engine is supercharged, but it does not have to supply continous power to a supercharger.

GMR Stirling Thermal Engine

Description — The requirements of the ideal Stirling cycle serve as a basis for establishing the prerequisite elements of an engine intended to follow this cycle. In order to pass through the desired cycle, the working medium within the engine must be heated, it must be cooled, it must store and remove energy from a regenerator, and it must be compressed and expanded at the proper times. These requirements then prescribe the five major components of the Stirling thermal engine — the engine heater and a means of keeping it hot; the regenerator; the engine cooler and a means of keeping it cool; a displacer piston to control the movement of the working fluid through the heater, regenerator, and cooler; and the power piston to compress and expand the gas.

The arrangement of these five major elements is shown in the schematic diagram of Fig. 3, and a photograph of the installed engine is shown as Fig. 4.

The heater consists of a bank of closely spaced stainless steel tubes brazed into the cylinder head which form continuous gas passages between the cylinder head and the regenerator. These heater tubes are kept hot by burning fuel in an external combustion system which is discussed below. The regenerator is a mass of fine wires sealed within eight individual cups located around the outside of the engine. The cooler is formed from bundles of small tubes surrounded by an annular chamber through which water is circulated. These cooler tubes form gas paths from the regenerator back to the cylinder.

The displacer piston is simply a hollow stainless-steel shell which fits loosely in the cylinder, and its motion is controlled through the displacer piston rod connecting it to the drive mechanism in the crankcase. There is very little pressure difference between the top and bottom of the displacer piston, and small leakage of gas between the displacer and the cylinder wall can be tolerated; so this piston requires no piston rings or other means of sealing to the cylinder. The top of the displacer piston operates approximately at the temperature of the heater while the bottom operates approximately at the temperature of the cooler, so the piston walls are made as thin as possible to reduce conduction heat losses, and the piston has two internal, horizontal baffles to reduce radiant heat losses through the piston.

The power piston must provide the work of compression of the gas and it must deliver the work of expansion, so it experiences considerable pressure differences between its upper and lower extremities. Therefore, it fits more closely into the cylinder and is equipped with suitable rings or other sealing means to prevent gas leakage between the piston and the cylinder wall. The power piston is entirely in the cold zone of the engine which allows more latitude in the selection of sealing means than do more usual engine pistons which experience very high combustion gas temperatures. Force is transmitted from the power piston to the rhombic drive mechanism by means of the power piston rod as shown. The power piston rod is hollow, and the displacer piston rod passes through it.

Photographs of the pistons are shown with the rhombic drive components in Figs. 11 and 12.

Figs. 3 and 4 both show a buffer space below the power piston and above a seal assembly at the top of the crankcase. The discussion of the inherent characteristics of the Stirling cycle showed that the engine would tend to have very low specific output if it were operated at low mean cycle pressures. Accordingly, the engine is designed to operate at higher pressures, and the buffer space is provided so that a thermodynamically inert volume of working fluid can be trapped below the power piston. The mean pressure of this gas in the buffer space is kept near the mean pressure of the working fluid above the power piston and serves to reduce the loads on the drive mechanism. The GMR engine operates with mean pressures of the working fluid near 1000 psi.

The seal assembly shown at the power piston rod seals this buffer pressure while the crankcase is vented and operates at atmospheric pressure. It is also evident that a seal is required between the displacer piston rod and the inside of the power piston. This seal is indicated in the drawing. Conventional "O" ring seals are used in both locations. A photograph of the inside of the power piston cylinder with power piston seal assembly is shown as Fig. 5. The buffer space is formed by the hollow space between the cylinder wall and the outside of the casting.

It has already been emphasized that the Stirling engine is an external combustion engine with the fuel and combustion products never entering the working spaces of the engine. A schematic diagram of the external combustion circuit is shown as Fig. 6. The essential elements of this external circuit are the preheater which heats the incoming combustion air with heat from the departing exhaust products, the combustion chamber where fuel is burned, and the engine heater tubes through which energy is transferred from the combustion products to the working gas in the engine cylinder. Sche-

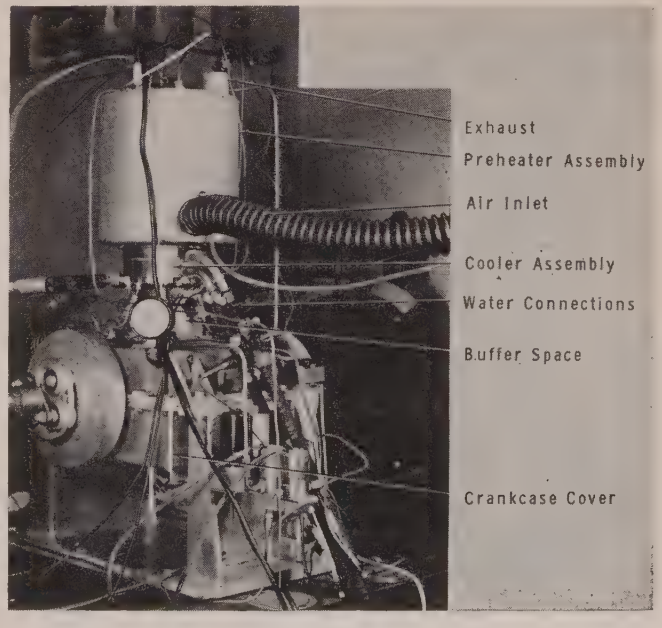


Fig. 4 — GMR Stirling thermal engine

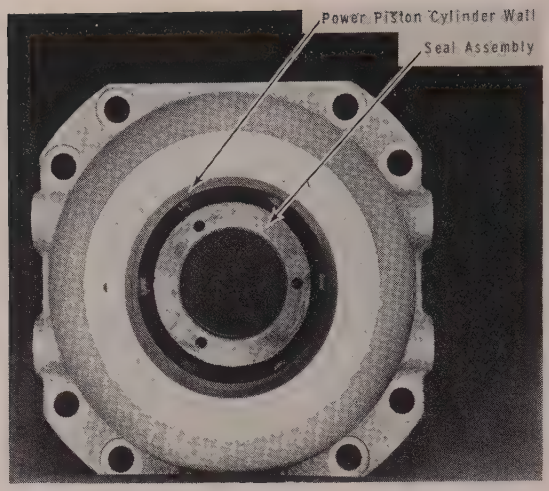


Fig. 5 — Power piston cylinder

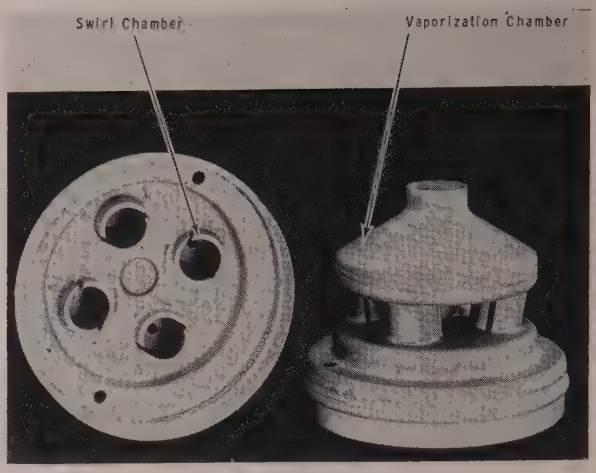


Fig. 8 — Stirling engine combustion chamber

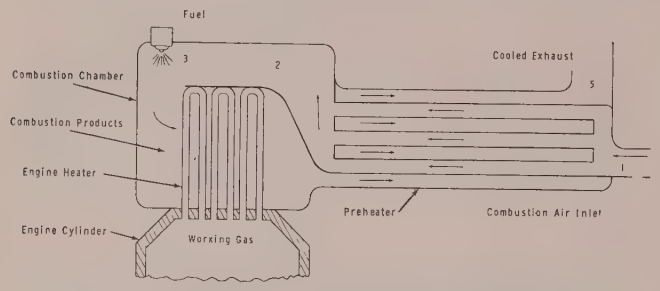


Fig. 6 — External combustion circuit schematic diagram

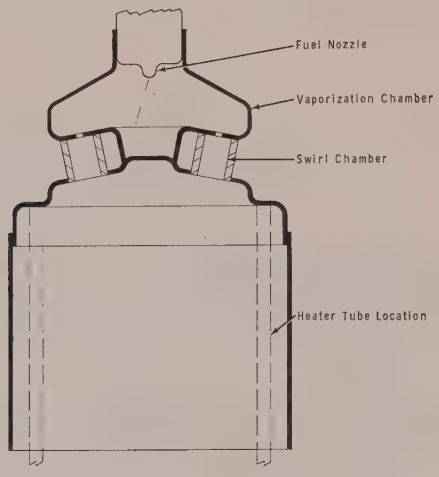


Fig. 7 — Stirling engine combustion chamber

matically, cold combustion air from a suitable blower enters at point 1. It is heated in the counterflow preheater, becoming hot combustion air at 2. Fuel is introduced by the nozzle and burned at 3, and the products of combustion give up energy to the working fluid in the heater tubes at 4. The exhaust gases then pass through the preheater where they are cooled by the incoming air and exhausted at 5.

In the actual engine construction, these components of the external combustion circuit are annular elements arranged symmetrically around the cylinder centerline. The counterflow preheater passages indicated in Fig. 3 are flat spiral passages which with their headers and ducts form an insulating assembly completely surrounding the hot parts of the engine. The exhaust gases flow outward through one set of spirals while the entering air passes inward through the alternate spiral passages. This preheated combustion air moves up through the headers and enters the combustion chamber. Fuel is introduced through an atomizing nozzle at the center top of the combustion chamber, and combustion takes place. The products of combustion pass radially outward and downward through the circular ring of closely spaced engine heater tubes and then out into the headers and spiral passages of the preheater. The cooled exhaust gas is then released through the exhaust outlet shown. The small numbers in the combustion gas circuit in Fig. 3 correspond to similar numbers in the schematic circuit of Fig. 6.

The combustion chamber used in present GMR Stirling thermal engines is more complicated than the one shown in Fig. 3. A cross-section of the actual chamber is shown as Fig. 7. The actual burner introduces the atomized fuel from the nozzle into a vaporization chamber where it is completely vaporized in an atmosphere much too rich to support combustion. This fuel vapor then passes into four swirl chambers, and it is mixed with the heated combustion air which enters the chambers through slots in their walls. Combustion is initiated in these swirl chambers. Fig. 8 shows the upper part of the combustion chamber with the vaporization chamber and the swirl chambers. This burner provides a short nonluminous flame with a high heat release rate in a small burner volume.

Fig. 9 shows the heater tubes and the main engine cylinder assembly to which they are brazed. The regenerator cups are also indicated. The two long tubes extending from the top of the heater assembly permit introduction of thermocouples, and the nut that clamps the cylinder assembly to the buffer space is also shown.

Rhombic Drive Mechanism — The motion of the pistons is controlled by the rhombic drive mechanism shown in Fig. 3. Fig. 10 is a view looking up into the crankcase with the crankcase cover removed, and it shows some of the parts which are not evident in the schematic diagram.

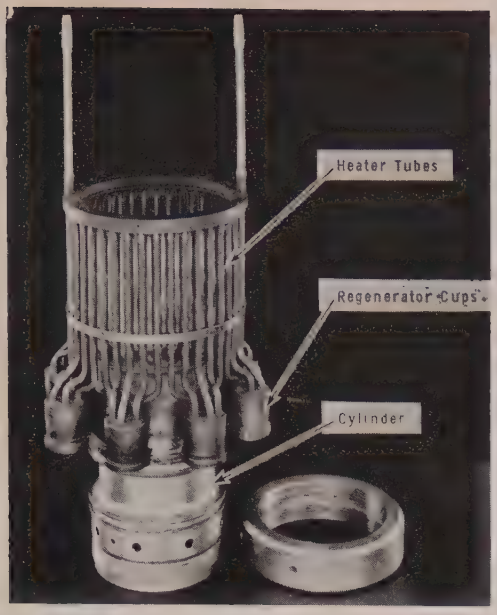


Fig. 9 — Stirling engine heater tube and cylinder assembly

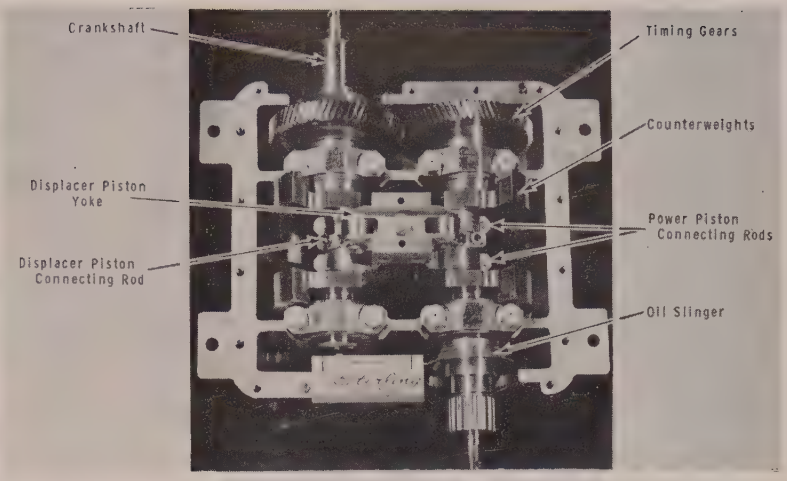


Fig. 10 — GMR engine rhombic drive

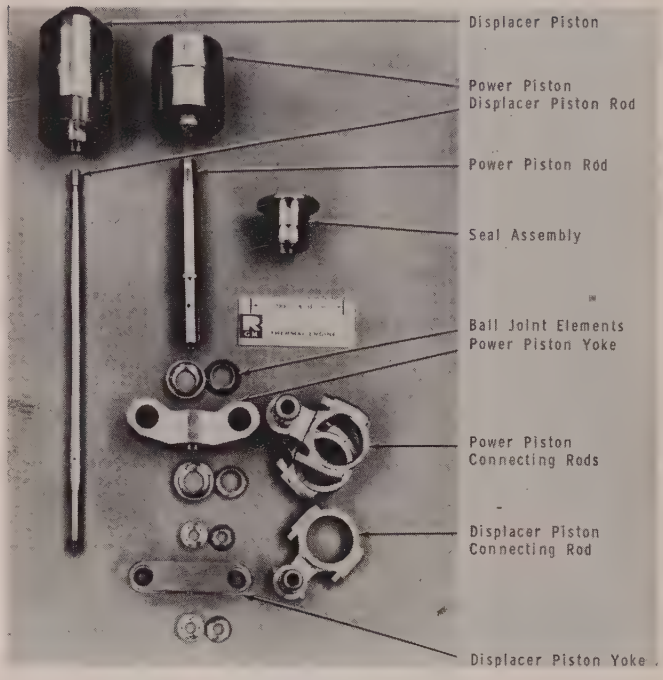


Fig. 11 — Rhombic drive components

The rhombic drive consists of two similar, contrarotating crankshafts with one crank on each crankshaft. These shafts are timed together by the pair of helical timing gears. The power piston is connected to the cranks by the power piston rod, the power piston yoke, and the power piston connecting rods. A similar series of parts connects the displacer piston to the crankshafts through the lower yoke and connecting rod assembly. There are actually two power piston connecting rods on each crank, and they are spaced by the yoke to serve as a single, split connecting rod. The lower, displacer piston connecting rod then fits on the crank between the two power piston connecting rods. This arrangement is clearly shown in the photograph of Fig. 10.

The individual components of the drive mechanism are shown in Fig. 11, and the drive components are shown in their assembled relationship in Fig. 12. The latter figure shows two positions of the drive. On the left the cranks are at their innermost points and the pistons are closest together, while on the right the cranks are at their outermost points and the pistons are farthest apart.

It is interesting to note that the pistons reach top dead center or bottom dead center when the crankshaft center, the crankpin center, and the yoke pin center are in a single straight line; and it is further to be noted that the positions of top dead center and bottom dead center of one piston are not 180 crank angle deg apart.

An important feature of this drive mechanism is its symmetry of construction which permits complete dynamic balancing of all the moving parts of the engine by means of suitably located and sized counterweights on the crankshafts. These counterweights are also indicated in Fig. 10.

Engine Operation — The operation of the Stirling thermal engine is completely governed by the relative motion of the two pistons, and it is most easily described by considering these piston functions.

The basic functions of the pistons may be seen

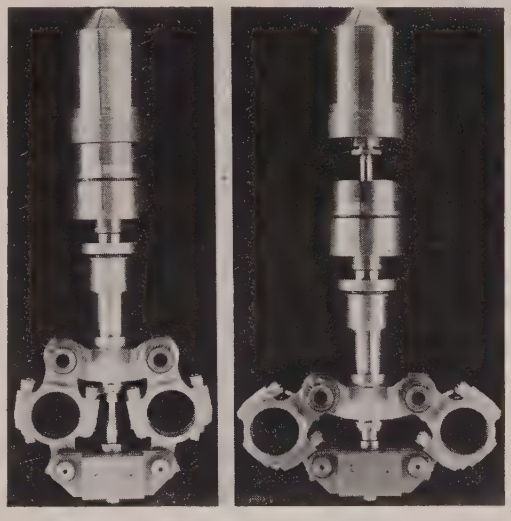


Fig. 12 — Rhombic drive subassembly

from an examination of Fig. 3. First, it is evident that all of the working fluid is contained in the total volume that is the sum of the volumes of the engine's five working spaces and their short connecting passages. These five spaces are: the variable volume of the hot space in the cylinder above the displacer piston, the internal volume of the heater, the gas volume of the regenerator, the internal volume of the cooler, and the variable volume of the cold space between the pistons in the cylinder. The total volume occupied by the working fluid is, therefore, completely governed by the position of the power piston. Since the displacer piston always occupies the same volume, its position cannot change the total volume of the system. Thus, gas volume change to effect its compression or expansion is accomplished by motion of the power piston and is independent of the position or movement of the displacer piston.

On the other hand, the displacer piston serves to control the location of the gas within the total volume established by the power piston. When the displacer piston is at the top of the cylinder, most of the gas is in the cold spaces. Downward movement of the displacer piston forces the gas through the heat exchanger circuit upward through the cooler, regenerator, and heater into the hot spaces. The net effect of this downward motion of the displacer piston is to move most of the gas up from the cold space to the hot space. Moving the displacer pistons down heats the gas. Reverse motion of the displacer piston upward from the bottom of its travel transfers most of the working fluid from the hot space downward through the heater, regenerator, and cooler into the cold space, and thereby effects a cooling of the gas.

Thus, the primary function of the power piston is to compress and expand the working fluid, and the primary function of the displacer piston is to heat and cool the working fluid. It must be emphasized that only the gas contained in the active spaces above the power piston goes through the thermodynamic cycle that converts part of the heat from the heater into work at the power piston. The gas contained in the buffer space below the power piston is the same kind of gas as the working fluid, but it accomplishes no thermodynamic purpose and serves only to balance the forces of the mean pressure of the supercharged working fluid.

The ideal Stirling cycle is usually described by the sequence of piston positions shown in Fig. 13A. The ideal PV diagram is repeated as Fig. 13B. In position I, the power piston is at bottom dead center (bdc), and the displacer piston is at top dead center (tdc). Thus, the gas is contained in the cold space at maximum total volume. The first process in the ideal cycle is an isothermal compression from I to II, and this could be accomplished by movement of the power piston from bdc to tdc without moving the displacer piston. The next required process from II to III is a constant volume heating of the gas. This could be accomplished by moving the displacer piston from tdc to bdc without moving the power piston. The third process is an isothermal expansion from III to IV, and this could be accomplished by moving the power piston from tdc to bdc without moving the displacer piston. The cycle would then be closed by returning the displacer piston to tdc to cool the gas at constant volume.

The above ideal description of the piston motion is based on the postulate that heat is transferred into or out of the working fluid through the cylinder walls during the isothermal processes. It has already been shown, however, that the actual configuration of the engine is based on the premise that heating and cooling of the working fluid must be accomplished by suitable movement of the displacer piston. Recognizing this inherent characteristic of the engine, a more realistic picture of the piston motion required to produce the ideal Stirling cycle can be created by means of a piston relative motion diagram like that of Fig. 13C.

In this diagram the position of the displacer pis-

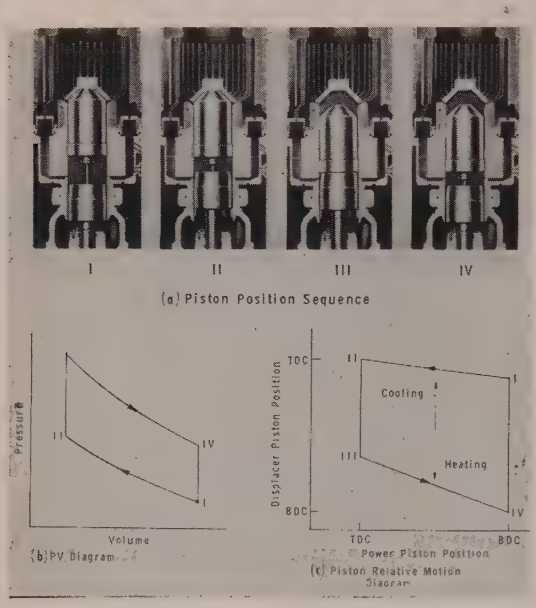


Fig. 13 — Ideal Stirling cycle diagrams

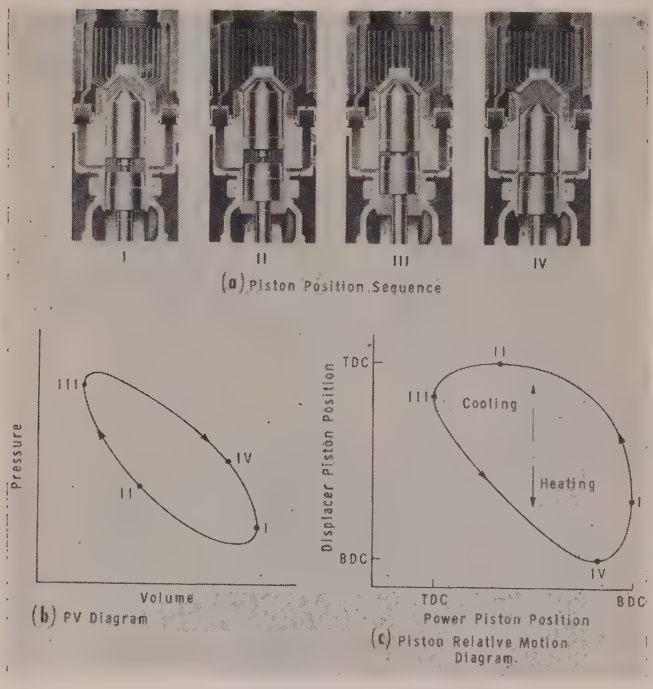


Fig. 14 — Actual Stirling cycle diagrams

ton is plotted on the ordinate as a function of the position of the power piston as the abscissa. The tdc's are at the top and left, respectively. This diagram can be used to illustrate the individual effects of the piston motions when the pistons move simultaneously.

It is possible to construct an ideal piston motion diagram from the ideal PV diagram of the cycle. The first, isothermal compression, process must be accomplished by movement of the power piston from bdc to tdc to reduce the volume of the fluid and by an upward movement of the displacer piston to provide cooling equivalent to the work of compression performed by the lower piston. Thus, the displacer piston cannot be at tdc at point I, but must rise from some lower position to tdc during the compression process. The second, constant volume heating, process from II to III can be accomplished with movement of only the displacer piston, but it cannot move the full stroke to bdc since heating must also be done during the isothermal expansion stroke from III to IV. After the constant volume heating process, the isothermal expansion is accomplished by moving the power piston from tdc to bdc while the displacer piston finishes its travel to bdc. The final process is the constant volume cooling from IV to I, and this may be accomplished by motion of the displacer alone. The complete piston motion diagram is then as shown in Fig. 13C for the ideal cycle.

It is impossible to draw this diagram to any exact energy scale, because the quantitative relationship between the displacer piston position and the amount of heat transferred to the working fluid cannot be expressed directly. Nor is it possible to deduce the correct shape of the curves between the corners of the diagram. It is, however, possible to make some approximate statements about the overall proportions of the diagram. First, the amount of heat transferred to and from the regenerators during the constant volume processes is much greater than the heat added during the expansion stroke or the heat removed during the compression stroke. Also, the heat added during the expansion must be greater, by the amount of work produced per cycle, than the heat removed during the compression stroke. It is, therefore, to be expected that the ideal movement of the displacer piston will be less between I and II than it is between III and IV, and that the movements between II and III and between IV and I will be greater than either of the two above.

The diagrams of Fig. 13 are duplicated in Fig. 14 using the actual engine parameters. A comparison of the actual PV diagram of Fig. 14B with the ideal diagram of Fig. 13B shows that the four ideal processes of the ideal cycle do not seem to be even closely approximated by the actual engine. However, a comparison of the relative motion diagrams (Figs. 14C and 13C) shows that the positions of piston top and bottom dead centers are clearly defined for both the actual and ideal engines. This correlation of cycle events and mechanical piston events allows comparison between the actual engine operation and the ideal cycle. The ideal cycle thus becomes a cycle of four piston events rather than of four processes as shown on ideal PV diagrams. The following four events follow in order for both ideal and actual engines: I — bdc power piston; II — tdc displacer piston; III — tdc power piston; IV — bdc displacer piston. These four events are easily noted on the piston motion diagrams, and they can be noted accurately on the actual PV diagram by equivalent crankshaft positions.

It is now possible to examine Fig. 14C and observe the actual engine piston movements and some of the compromises that are imposed by the real drive mechanism. The first compromise which is evident results from the impracticability of allowing the power piston to completely stop for part of the cycle. It is theoretically possible to conceive a drive which would allow this discontinuity, but any reasonable mechanism will not permit it. Thus, the vertical lines of Fig. 13C have been eliminated. It would also be impossible to stop the displacer piston during part of the cycle, but this is not required even in the ideal instance as shown by the absence of horizontal lines in Fig. 13C.

Beginning at point I, the actual engine events can be traced on Fig. 14C. It is desired to accomplish isothermal compression from event I to event II, and the figure shows that the displacer piston moves in the cooling direction quickly at the beginning of this compression so that compression occurs with most of the fluid being cooled in the cold spaces, but cooling stops before the power piston reaches tdc. It is seen that the process from II to III is heating as in the ideal cycle, but that it is accomplished at decreasing gas volume rather than at the ideal constant volume. The expansion from III to IV, however, is quite similar to the ideal. The pistons are moving down together as indicated by the nearly straight, 45 deg, portion of the curve; and the downward motion of the displacer tends to keep the gas in the hot spaces where it is heated while the downward motion of the power piston accomplishes the expansion or power stroke. The final, cooling process is accomplished by an upward movement of the displacer piston while the power piston completes its travel to bdc.

The actual engine operations during the expansion and compression strokes are quite like those anticipated from the ideal considerations with a relatively large time for heating during the expansion (III–IV) and a relatively small time for cooling during the compression (I-II). However, the long vertical lines which characterized the ideal diagram during the constant volume, regenerative processes are not in evidence on the actual diagram. This difference is easily explained by examining the heat exchanger circuit in Fig. 3.

The figure shows the heater, regenerator, and cooler in a single continuous passage, so that gas passing from the cold space to the hot space, for example, must pass through all three heat exchangers. It has been stated that the net effect of moving gas from the cold space to the hot space is to heat the gas. The truth of this statement is evident, and it is an adequate approximation of the major cycle event accomplished by downward motion of the displacer piston. It will be seen, however, that some of the heat introduced into the working fluid during its upward movement is recovered from the regenerator since the cold gas passes through the regenerator before it enters the heater. The converse of this heat transfer takes

place during downward motion of the gas through the heat exchanger circuit when heat from the heater is stored in the regenerator before the gas passes to the cooler and the cold space of the cylinder.

The difference between the actual relative motion diagram and the ideal relative motion diagram (Figs. 14C and 13C) can thus be explained. In the ideal diagram, the regenerative processes were clearly shown, because it was presumed possible to stop the power piston while these constant volume heat exchanges took place. In the actual diagram the vertical lines are absent, because the power piston does not stop; but the regenerative process still takes place during the entire upward or downward motion of the displacer piston while the net cooling or heating is being accomplished.

The above discussion suffices as an introduction to the means by which the Stirling engine performs the heat transfer processes necessary to convert part of the heat from the heater into work at the power piston. The explanation is by no means complete, and it overlooks some significant effects; but it illustrates how the proper thermodynamic events can take place within the engine.

The discussion of the heat flows is sufficient to demonstrate how the work is done, but one further statement might be made about the pressure variations of the working fluid. Except for flow losses through the heat exchangers and kinetic effects, the static pressure of the working fluid is constant in the engine working spaces at any given moment. The magnitude of this pressure is determined by the gas law relationships and the distribution of the mass of working fluid in the various spaces of the engine. The pressure of the gas in the system is influenced by motions of both pistons. If the displacer would move the gas from hot spaces to cold spaces with no motion of the power piston, the gas pressure would decrease. Also, the power piston could change the gas pressure with no motion of the displacer piston by changing the total volume of the system. With the actual simultaneous motion of both pistons a combination of these basic effects is produced.

The power output of the Stirling thermal engine is regulated by changing the mass of working fluid passing through the thermodynamic cycle. The cycle itself is not changed. A fuel control keeps the heater at a constant temperature regardless of the engine load level, and the cooler remains at a temperature near that of the available cooling water. The engine thus always produces a constant amount of work per pound of working gas, and the output is changed by changing the mass of working fluid in the active engine spaces. In the present GMR laboratory engines, this control is accomplished by supplying hydrogen from a high-pressure bottle, or bleeding working fluid to the atmosphere. An automatic scheme for providing regulation has been devised by Philips, and it is discussed by Meijer.[1]

A unique characteristic of the Stirling thermal engine and one which tends to indicate the degree of perfection with which it performs the thermodynamic conversion of heat to work is the capability of the engine to function as a heat pump when work is put into the crankshaft instead of being taken from it as in normal engine operation. If the fuel and the burner are turned off while the engine is producing power, the output of the engine will gradually decrease as the heater is allowed to cool. At some point the heater temperature will become so low that the engine will no longer supply enough power to overcome its own friction, and the engine normally will stop. However, if the engine is maintained in rotation in the same direction by driving the crankshaft from an external machine, the Stirling thermal engine will become a heat pump and continue to pump heat from the heater to the cooler. A few moments of this operation will cool the heater pipes sufficiently to cause frost to form from the moisture in the air, and the engine will continue to refrigerate the surroundings as long as power is supplied to the crankshaft.

If the engine is motored with rotation opposite to that with which it produces power as an engine, it will also act as a heat pump; but the direction of heat flow will reverse with the reverse direction of rotation. Under these conditions, the engine will pump heat from the cooler to the heater; and a few moments of this operation will cause the heater tubes to become hot and heat the surroundings.

This heat pump or refrigerator operation of the engine is easily explained from the principles of thermodynamics, and it is the performance that is expected as a matter of routine from reversed thermodynamic cycles; but it is rare, indeed, to find an engine that understands its thermodynamics as well as the Stirling thermal engine!

Table 2 — Large Engine Comparison

Engine Ideal Cycle Fuel	Stirling Stirling Diesel No. 1	6-cyl diesel 2-stroke diesel Diesel No. 1	Automotive V-8 Otto Gasoline
Fuel Heating Value, Btu/lb	18,200	18,200	18,900
Bore, in.	3.47	4.25	4.06
Stroke, in.	2.37	5.00	3.56
Displacement/Cylinder, cu in.	20.08	71	46.25
Total Engine Weight, lb	450	2190	678
Maximum Economy Data			
Minimum Bsfc, lb/bhp-hr	0.358	0.40	0.415
Brake Horsepower	30	181	199
Speed, rpm	1500	1700	3000
Max. Thermal Efficiency, %	39.0	34.8	32.4
Full-Load Data			
Total Horsepower	40	210	242
Hp/Cylinder	40	35	30.2
Speed, rpm	2500	2100	4600
Bsfc, lb/bhp-hr	0.418	0.41	0.468
Thermal Efficiency, %	33.3	34.0	28.7
Bmep, psi	317	93	113
Specific Weight, lb/hp	11	10.4	2.80
Specific Output, hp/cu in.	2.00	0.49	0.65

Modern Stirling Thermal Engines

Performance data for a relatively large Stirling engine are compared with other large engines in Table 2, and a similar comparison for smaller engines is given in Table 3.

The Stirling engine data in the two tables are taken from General Motors Research tests on the original two engines brought to this country by Philips. The diesel engine data are taken from

[1] "Philips Hot Gas Engine with Rhombic Drive Mechanism," by R. J. Meijer. *Philips Technical Review*, Vol. 20, 1959, pp. 245–262.

Table 3 — Small Engine Comparison

Engine	Stirling	4-cyl otto	6-cyl outboard
Ideal Cycle	Stirling	4 stroke	2-stroke otto
Fuel	Diesel No. 1	Isooctane	Gasoline & oil mixture
Fuel Heating Value, Btu/lb	18,200	19,160	18,900
Bore, in.	2.36	2.50	2.44
Stroke, in.	1.33	2.25	2.13
Displacement/Cylinder, cu in.	4.94	11.0	10.0
Total Engine Weight, lb	127	120	96
Maximum Economy Data			
Minimum Bsfc, lb/bhp-hr	0.47	0.458	0.81
Brake Horsepower	5.98	21	29
Speed, rpm	2400	3600	3500
Max. Thermal Efficiency, %	29.6	29.0	16.6
Full-Load Data			
Total Horsepower	8.63	25.6	48.5
Hp/Cylinder	8.63	6.4	8.09
Speed, rpm	3600	4000	5200
Bsfc, lb/bhp-hr	0.528	0.51	1.10
Thermal Efficiency, %	26.4	26.1	12.2
Bmep, psi	192	115	61.5
Specific Weinght, lb/hp	14.7	4.68	1.98
Specific Output, hp/cu in.	1.74	0.58	0.81

Hulsing and Ervin.[2] The data on the remaining engines are taken from GMR tests on commercial engines of the type described. Performances of the engines at maximum economy and at full load are considered in each instance. The large engines all produce 30–40 hp per cylinder at full load, while the small engines are in the 6–9 hp per cylinder group. All of the engines except the Stirling engines are multiple cylinder engines and, therefore, are larger in total output.

In Table 2 it is seen that the maximum thermal efficiency of the Stirling engine is greater than that of either of the other two engines, but that the efficiency of the Stirling engine falls slightly below that of the diesel engine at full-load conditions. The automotive engine selected for this comparison produced thermal efficiencies about 4% greater than the average of all similar engines for which data were available. The diesel engine is one specifically designed to give maximum economy of operation.

The full-load data shows the Stirling engine to be slightly heavier, per horsepower, than the diesel, and that both of these are considerably heavier than the automotive engine.

The Stirling engine is thus seen to be similar to the diesel engine in efficiency and specific weight; and it is more efficient, but heavier than an automotive engine.

The small engine data of Table 3 again show the Stirling engine to have the highest maximum thermal efficiency as well as the highest efficiency at full load. In these small engine comparisons, the Stirling engine is seen to be heavier than either of the other engines.

Both of the tables give the outputs of the various engines per cubic inch of piston displacement, and the Stirling engine is markedly superior to any of the others according to this criterion. This volume rating is probably not a suitable way to compare the effectiveness of the use of the active spaces of the engines, because the power piston swept volume is nearly the entire active volume of the internal combustion engines while it represents only about one-third of the total active space of the Stirling engines.

The mep data, however, do indicate the high specific performance of the Stirling engines which results from the supercharging effect of the high pressure working fluid. The mean pressure of the working fluid in the small Stirling engine is 1000 psi and that of the large engine is 1500 psi. A further comparison of engine pressures is given in Fig. 15 which gives pressure-crank angle curves for the small GMR Stirling thermal engine and two-stroke diesel engine. The Stirling engine buffer space pressure is also shown. It will be noted that the pressure variations of the Stirling engine have more gradual rates of change and are less extreme in their total travel than are the pressure variations of the diesel engine.

It might be well to consider the effect of comparing single-cylinder Stirling engines against multicylinder otto and diesel engines. In the larger sizes, there is probably not much effect on the performance parameters shown in Tables 2 and 3 except a slight weight disadvantage for the single-cylinder engine.

In the smaller sizes parasitic effects are more significant and the total output of the engines should be considered. It will be seen that the 8.6-hp Stirling engine has better thermal efficiency than the highly refined, 4-stroke cycle, 25-hp engine, and more than double the efficiency of the 2-stroke, 48-hp engine. Experience would indicate that if either of the latter two engines were scaled down to the same total output as the Stirling engine that they would suffer a loss in efficiency from that of the table. The 40-hp Stirling engine of Table 2 is much more efficient than the larger power otto engines of Table 3.

In the small engines, as in the large engines, the single-cylinder engine suffers a weight disadvantage. It is difficult to judge the extra weight of the small Stirling engine quantitatively from Tables 2 and 3, however; because the Stirling engines shown are laboratory test engines which were constructed with no attention to weight reduction, while all of the other engines are commercial products of sophisticated design from which all excess weight has been carefully removed.

The remaining, and important, advantages of the Stirling engines do not lend themselves to numerical tabulations as do the performance parameters. Along with its inherently high efficiency, the comparatively silent operation of the Stirling engine is its greatest asset. The description of the rhombic drive stated that the mechanism of the engine could be completely balanced for one or more cylinders at any speed. This characteristic is indeed unique to the Stirling engine. It is also evident from the description of the engine that there are no valves with their attendant valve train noises, and the pressure variation curves of Fig. 15 show no sharp pressure pulses to excite noise and vibration.

A further indication of the forces involved in the drive of the Stirling engine is given by Figs. 16 and

[2] "GM Diesel's Additional Engines — New Vees and In-Lines," by K. L. Hulsing and C. E. Ervin. SAE Transactions, Vol. 67, 1959, pp. 596–618.

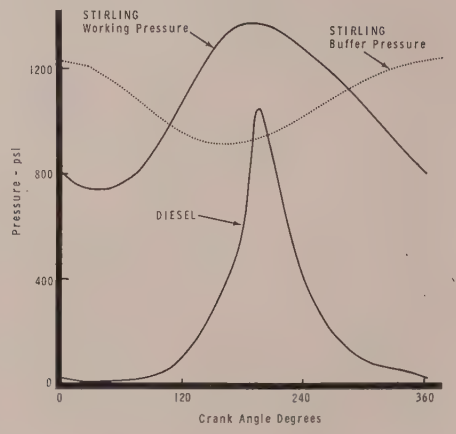


Fig. 15 — Pressure-crank angle curves

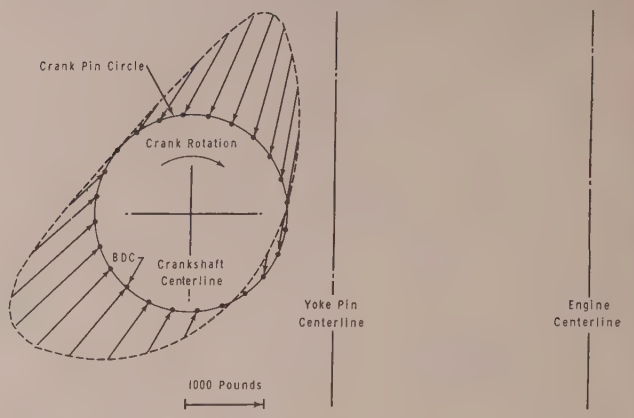


Fig. 16 — GMR Stirling engine crank pin forces

17. Fig. 16 is a scale drawing of some of the important parameters of the GMR engine rhombic drive. The engine centerline, and the vertical centerline of the power piston yoke pin are indicated. The center of the crank of the left hand crankshaft travels around the circle shown, and vectors indicating the direction and magnitudes of the power piston connecting rod force on the crank are shown. Fig. 17 shows the variation in torque on one crankshaft as a function of crank angle. The zero of the crank angles is taken at bottom dead center of the power piston for Figs. 15–17.

When Stirling engines are operating, the most evident noise is that of the timing gears on the crankshafts. The larger Stirling engine sounds very much like an electric motor when operating under load, while the smaller engine is somewhat noisier due to the higher operating speed.

The capability of the Stirling external combustion engine to operate from a wide range of fuels or from another type of heat source has also been made evident; and in special circumstances, this characteristic of the Stirling engine would prove to be of significant advantage. On the other hand, the Laboratory Stirling engines produce cool, colorless, odorless, carbon-monoxide-free exhaust when operating on No. 1 diesel fuel.

Some of the disadvantages of the Stirling engine become apparent when the complete powerplant and its accessories are considered. First, most of the heat rejected from the Stirling cycle is rejected to the cooler with very little leaving to the atmosphere through the exhaust gases. Thus, the Stirling engine requires a larger radiator than the internal combustion engine, but the high efficiency of the Stirling engine tends to minimize this effect. It is estimated that a Stirling engine would require a radiator about 2.5 times as large as that of an automotive engine for the same shaft horsepower.

The discussion of the external combustion system also points out a source of relative demerit for the Stirling engine in that it requires a blower to force the air through the preheater and the combustion chamber. This blower constitutes a parasitic power loss for the engine and could well be a significant source of noise. Here again, the high basic efficiency of the engine allows some compromise. In circumstances where the ultimate maximum efficiency is not required, the blower noise and power can be diminished by reducing the flow loss and the effectiveness of the combustion air preheater with an attendant loss in overall engine economy.

Some comment about the relative costs of the engine is also in order, even though the Stirling engine is still far from the production stage at which comparative dollar values could be established. The Stirling engine will cost appreciably more than the other modern engines, primarily because of the complexity of the heat exchanger construction required. The heater is constructed of small stainless steel tubes which must be carefully spaced to achieve the optimum heat transfer configuration and which must be carefully brazed to form hydrogen-tight joints under high temperature operating conditions. The cooler must likewise be hydrogen-tight, but it operates at very low temperatures, and is therefore less costly to construct. The complete balance and careful piston timing achieved by the rhombic drive are obtained at the expense of considerably more mechanical complexity than that of a simple crank mechanism. It is anticipated that one of the results of the new GMR research program will be the elimination of some of this cost with a minimum sacrifice in engine performance.

A brief, objective, statement of the place which the modern Stirling engine will take among the prime movers of the future would seem to be as follows:

For large engine installations, the weight and efficiency will be at least as good as those of high-economy diesel engines; and the Stirling engine would be used where quiet operation, invisible, odorless exhaust or use of special fuels or heat sources warranted additional cost.

In small engine installations of 10 hp or less, the Stirling engine would be used where its significant efficiency and noise advantages compensate for extra weight and cost. The special fuel capability and the clear exhaust would serve as advantages here also.

The aforementioned advantages of the Stirling thermal engine make it particularly attractive for certain military applications, and GMR is currently building, under contract with the U. S. Army Engineer Research and Development Laboratories, a 3-kw ground power unit that will drive a 60-cycle electric generator for some military application. This portable, silent ground power unit will con-

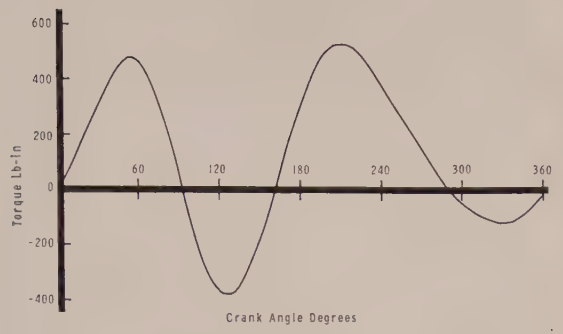


Fig. 17 — GMR Stirling engine crankshaft torques

THE ENGINEER.

ANTI-EXPLOSIVE STEAM BOILERS,

BY DUNN, HATTERSLEY, AND CO.,

MANUFACTURERS AND PATENTEES,

WINDSOR BRIDGE IRON WORKS, NEAR MANCHESTER.

The Engineers and the Public generally, have often asked for Non-Explosive

Fig. 18 — 1856 advertisement

sist of a Stirling thermal engine, the engine cooling system, and all necessary operating auxiliaries including the fuel tank and controls.

Historical Milestones

The principal factor which motivated 19th century inventors into efforts to build practical air engines seems to have been a widespread fear of steam boiler explosions. Fig. 18, from an 1856 engraving, illustrates the efforts some manufacturers took to allay these doubts. Early 19th century statistics are lacking, but in the years from 1862 to 1879 there were over 10,000 boiler explosions in England; and in the 40 years from 1880 to 1919 there were 14,281 recorded boiler explosions in the United States resulting in over 10,000 deaths and 17,000 injuries.[3]

Air Engines Defined — By the term "air engines" we refer to reciprocating heat engines which compress cool air and expand heated air. Purely rotating or turbine machines are excluded as is the internal combustion engine as we know it today, because historic air engines did not employ combustion directly in the cylinder. Air engine working fluids may be gases other than air, providing they do not change phase during the cycle. Air engines have been generally divided into three types which will be discussed in the chronological order in which they appeared. Type I engines were open cycle engines in which the products of combustion mixed with the working air passing through the cylinder. The cycle was similar to the Brayton cycle. Type II air engines were closed cycle, externally heated, and are called Stirling or constant volume engines. Type III were open cycle, externally heated, and generally know as Ericsson or constant pressure engines.

Types I and III will be discussed only briefly as they are thoroughly covered by Finkelstein.[4]

Superstition and Thermodynamics — It is difficult to review the history of the air engine without including the history of thermodynamics, for it was during the first half of the 19th century that the air engine was introduced and the understanding of the nature of heat was developed. As a matter of fact, the failure of the air engine to reach the same measure of success as the steam engine can be ascribed to a misunderstanding of its operating principles which resulted from the lack of understanding of the nature of heat. Although Carnot wrote his *Reflections on the Motive Power of Heat* in 1824, acceptance of the kinetic (dynamical as it was also known) theory of heat was not general until after the middle of the 19th century. The first experimental determinations of the mechanical equivalent of heat were made by J. P. Joule and published in England in 1843. Until this time, the caloric theory of heat was prevalent. According to this theory, heat absorbed in the expansion of a gas became latent, but remained in the gas and was again evolved on compressing the gas. The caloric theory gave no explanation of the source of work produced by gas expanding against a piston. Heat or caloric (Webster's, 1957 — "a subtle imponderable fluid") was considered a substance instead of being associated with energy as in the kinetic theory. This error resulted in many attempts to build perpetual motion heat engines during the 19th century. The famous British inventor, Sir William Siemens, stated in a letter in 1875 that when he presented his paper on the *Conversion of Heat Into Mechanical Effect* at a meeting on "Heated Air Engines" in 1853, "the believers in the dynamical theory of heat, might be counted on your fingers."[5]

Air Engines — Type I — Type I engines are usually referred to as "furnace gas engines," and they approximated the Brayton cycle in operation.

The first known air engine patent was by the Rev. Henry Wood, an English vicar, in 1759. He claimed a method of operating an atmospheric engine of the Newcomen type with air heated by, "passing either through a fire, or through red-hot pipes, or through boiling water."[6]

The invention of the first working furnace gas engine is credited to a prolific engineer, Sir George Cayley, and was described in *Nicholson's Journal* in 1807. Solid fuel was burned in an airtight combustion chamber near the expansion cylinder. Compressed air from an adjacent pumping cylinder was passed through the fuel bed and the products of combustion admitted to the power cylinder

[3] Courtesy of Hartford Boiler and Insurance Co.
[4] "Air Engines," by T. Finkelstein. *The Engineer*, Vol. 207, 1959, pp. 492–497, 522–527, 568–571, 720–723.
[5] "Air Engines." *Engineering*, Vol. 19, 1875, pp. 200–201, 241–242, 287–289, 320, 355–356, 363, 417–418, 504–505.
[6] "Inventor of Hot Air Engine." *The Engineer*, Vol. 186, 1948, pp. 168–169.

through a valve.

In the United States, Philander Shaw and S. H. Roper received numerous patents and built some furnace gas engines during the latter half of the 19th century. While the efficiency of some was reported to be over 7%, which exceeded the small engines of that day, they all suffered from the problem of high wear as a result of dust and ashes being carried directly into the expansion cylinder. Fig. 19 is a Roper coal burning engine on exhibit at the Henry Ford Museum, Dearborn, Mich.

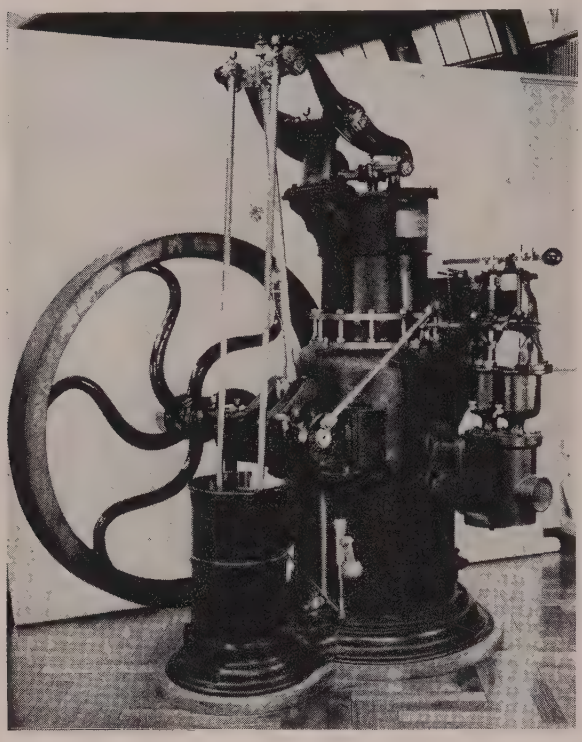


Fig. 19 — Roper coal-burning engine

Air Engines — Type II — Type II engines were closed cycle, externally heated engines which operated on what has become known as the Stirling cycle. Robert Stirling was born in Perthshire, Scotland, in 1790. He was licensed by the presbytery of Dumbarton in 1814, and was ordained on September 19, 1816. Less than two months later, he took out a patent, No. 4081, for a regenerator and an air engine which incorporated the regenerator. This was apparently the first suggested use of the regenerator in industry, Stirling specifically mentioning glass manufacturing, breweries, distilleries, and dye works. It was not readily apparent from the patent title that it dealt with air engines, since Stirling's main interest appeared to be the economizing of fuel. The title was, "Improvements for Diminishing the Consumption of Fuel, and in Particular an Engine Capable of Being Applied to the Moving of Machinery on a Principle Entirely New." That it was entirely new and that it was the work of a genius there is little doubt. The original patent drawings of Stirling's engine are reproduced as Fig. 20. An engine of this type was first operated in 1818.

Stirling's grandfather, Michael Stirling of Dunblaine, invented the first rotary threshing machine in 1756. Stirling's brother, James, was a member of the Institution of Civil Engineers and cooperated with him in construction and improvements of air engines during the years from 1818 to 1847. Four sons attained distinction in the profession of engineering, in such places as Scotland, Peru, and Hawaii.

Robert Stirling received the honorary degree of Doctor of Divinity from the University of St. Andrews in 1840. Apparently, he was greatly esteemed as a minister and was noted as a classical scholar

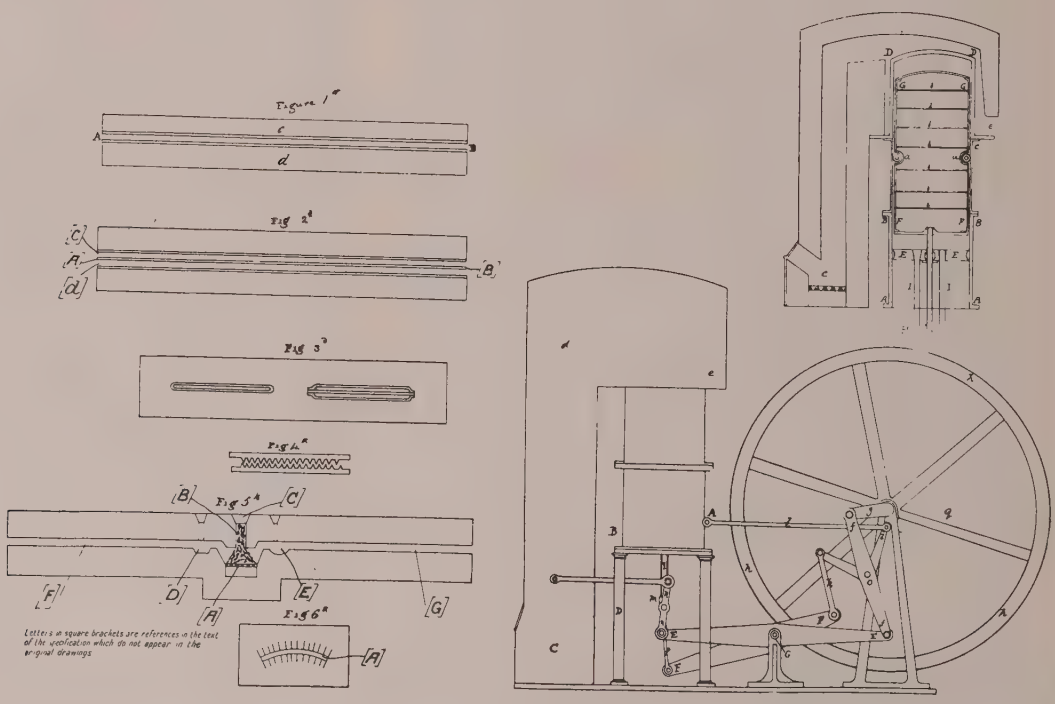


Fig. 20 — Stirling's patent drawings, 1816

as well as an inventor before his death in 1878.

A thorough explanation of the Stirling cycle has been given in the previous sections. It is of interest, however, to note several aspects of the original invention which were unique compared to other types of reciprocating engines. The Stirling engine had no valves nor any ports uncovered by pistons. It had the distinct advantage of locating the power piston in the cold zone, so that the piston remained cool throughout the cyclic process whether the engine was idling or at full load. Another advantage was its closed operation, the same working gas being used over and over again. This permitted the minimum cyclic pressure to be raised considerably above atmospheric, increasing the specific output proportionally. This feature was due to James Stirling, and the result was the brothers' most successful engine, which they installed in 1843 at the Dundee, Scotland, foundry. The mean effective pressure of this engine exceeded that of any subsequent air engine for the next 95 years.

Many commercial air engines produced during the latter part of the 19th century eliminated the regenerator which essentially is the engine's most important element. Apparently, this was done in order to simplify the design and decrease the cost, but this particular design simplification sacrificed the high potential efficiency of the Stirling cycle.

It was unfortunate that the true theory of heat was not better developed at the time Robert Stirling and his brother were building their engines. Although they recognized the vital contribution of the regenerator, the brothers believed that the heating surface was only required to make up the unavoidable losses due to radiation and friction. Robert Stirling, in fact, states in reference to his regenerator, "Nay, it is evident that by multiplying the members of the series indefinitely, air could be heated and expanded and made to do work at no appreciable expense. But let no mathematician be alarmed with the idea of a perpetual motion for the creation of power. There are many enemies to contend with the air engine besides friction, which alone prevents perpetuity in some mechanical motions." He then added, "the greatest enemy of the economizing principle, is the continual passage of caloric from the hot to cold parts of the engine, by radiation, conduction, and etc., which requires a continual supply of caloric to maintain the proper temperature of each."[7] In a paper presented before the Institution of Civil Engineers in 1845, James Stirling also expressed the belief that the engine could have been a perfect engine except for friction and heat losses. He said, ". . . . the hope of the perfect success of the engine, and particularly of its successful application to the propelling of ships, is principally founded on the fact that the powers of the economical process, which theoretically appear to be infinite, are not yet practically nearly exhausted."[8]

Because of these erroneous views, the Stirlings designed the heating surfaces to include only the area of the outer end of the displacer cylinder and its hemispherical head, with no provision for fins or projections to increase heat transfer surface. With insufficient area for the power level, the safe working temperature of the cast-iron parts exposed to the flame was exceeded and they usually failed within less than a year of operation. The son, Patrick Stirling, reported in 1860 that the heaters were, "kept red hot."[9] After four such heaters had cracked, the Dundee foundry gave up the use of Stirling's last air engine and replaced it with a steam engine in 1847. Materials were available at that time which might have enabled the Stirling brothers to increase the heater area by means of tubes as they had already successfully done with their rather sophisticated cooler designs described in the Stirling patent of 1840.

Improvements on the original engine design were patented jointly by the brothers in 1827 and again in 1840. Before building their final engine, a test model with a cylinder diameter of 3⅝ in. and a

[7] *Proceedings of the Institution of Civil Engineers* (England), Vol. 12, 1852.
[8] *Proceedings of the Institution of Civil Engineers* (England), Vol. 4, 1845, pp. 348–361.
[9] *Proceedings of the Institution of Engineers of Scotland,* Vol. 4, 1860, pp. 13–15, 40–50, 141–158.

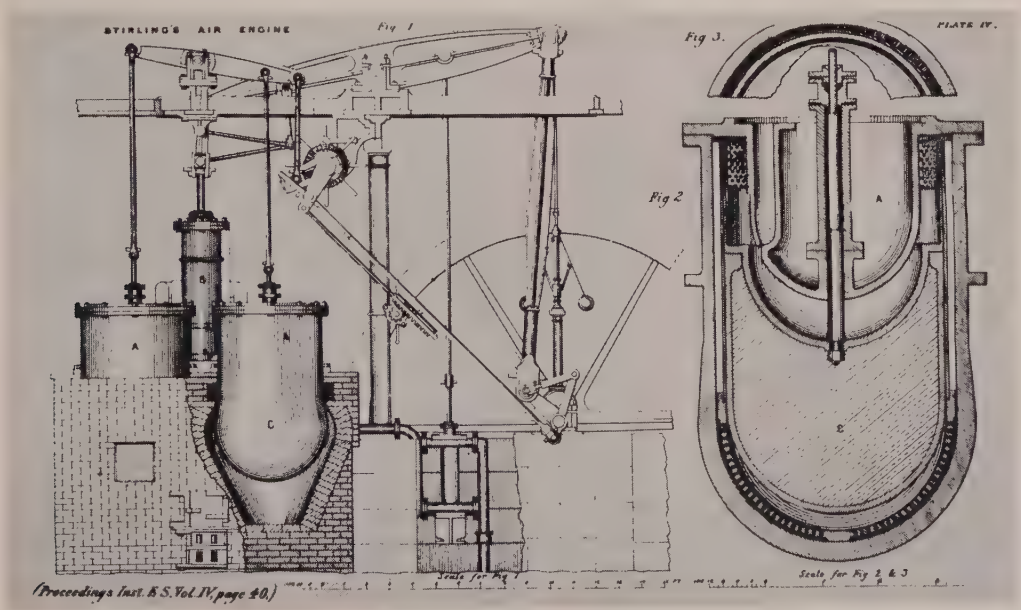


Fig. 21 — Stirling's last engine, 1843–1847

12-in. stroke was constructed and produced 2 hp. Their final engine (Fig. 21) developed nominally 45 hp and ran the lathes and other machinery at the Dundee foundry from March, 1843, to January, 1847. The power cylinder inside diameter was 16 in. and the piston operated with a 4-ft stroke at 28 cpm. By means of a prony brake they measured an average of 21 bhp, while the engine required 50 lb of coal per hour, or about 2.4 lb per bhp-hr. The best steam engines of the day consumed about 4 lb per bhp-hr. According to Stirling the coal was of inferior quality; but if we assume an "as received" higher heating value of 12,000 btu/lb, the brake thermal efficiency would have been 8.9%. Then, assuming a reasonable mechanical efficiency of 75%, and from additional data furnished by James Stirling on the heat to the cooling water, the indicated efficiency appears to have been 21% and the furnace efficiency about 57%. The temperature limits of the air were said to be 150 F and 650 F, but no measurements were known to have been made except of the cooler metal temperature. Carnot efficiency between these temperature limits is 45%.

As can be seen from Fig. 21, there were two displacer cylinders feeding one double-acting power cylinder. Each of the 4-ft diameter displacer cylinders contained an annual regenerator composed of several thousand $38 \times 1\frac{3}{4}$ in. sheets of 1/50 in. thick iron, with a heat transfer area of about 3200 sq ft. The minimum cyclic pressure at full load was about 150 psig, and the maximum about 240 psig.

In their patent of 1827 the brothers anticipated the advantages of using a working fluid other than air and apparently had tested other gases. They stated, "and though we generally work the engine with common atmospheric, yet whenever a supply can be conveniently obtained of nitrogen, carbonic acid (CO_2) or any permanent gas which does not readily corrode iron nor cause explosion, we use it with equal and in some respects greater advantage."

The engineers of the day were divided into three groups in regard to the utility of the regenerator, and it was a subject repeatedly debated in the British Institution of Civil Engineers. One group denied that the regenerator had any effect whatever. A second group, including the Stirling brothers and Sir George Cayley, concluded that a perfect regenerator would result in a perfect engine with practically no fuel consumption. The third group took the middle course: the regenerator was a useful device, but as to exactly how it worked they were not sure, although they were convinced that it could not produce a perpetual motion machine. There were apparently large groups of engineers who believed that compressed air was reduced in temperature after escaping through an orifice, and that expansion of air in a cylinder did not reduce its temperature. With these erroneous beliefs it is somewhat remarkable that the engines worked as well as they did.

The first published scientific explanations of the economizer, or regenerator, were by Siemens in 1853 and by Professor Rankine in 1854. Rankine, Carnot, and Joule are considered by many as the founders of the science of thermodynamics. Rankine, and J. R. Napier built a so-called improved air engine in the early 1850's (patented 1854) in which

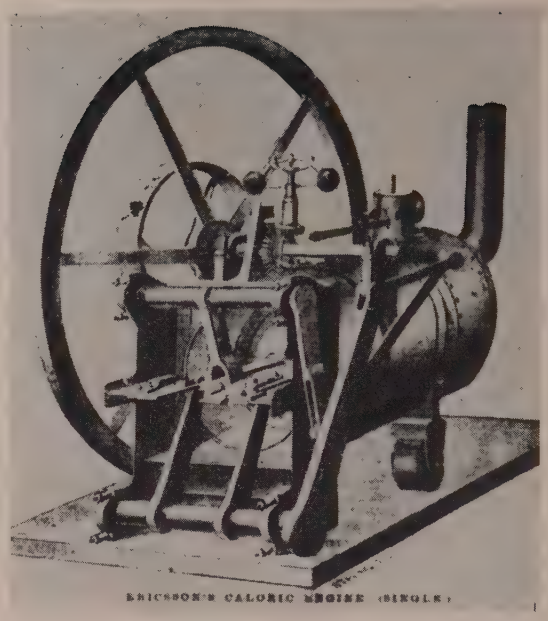


Fig. 22 — Ericsson's 1859 engine

they apparently had increased the heating surface by means of tubes which could be moved into and out of the hot air circuit during portions of the cycle.

Possibly because university level training for engineers was late in coming to Great Britain, there appeared to be a great lack of appreciation of the first and second laws of thermodynamics. At an engineering meeting in December, 1860, at which Patrick Stirling gave a paper on his father's engine developments, it is evident from the comments that among the 11 discussers only one, Professor Rankine, understood that the engine had to reject the difference between the heat input and the work output.

Even in 1875 there was still a great deal of controversy as to how the cycle operated. It was generally believed that the loss of efficiency was chiefly due to the heaters being constantly hot throughout the cycle, which was not true, as well as to inperfections of the regenerator. Rankine estimated the effectiveness of Stirling's 1845 engine regenerator at between 90 and 95%. This indicates that the major problem was not in the regenerator. Rankine had great respect for the Stirling cycle and stated that, "this was the best air engine that had been put in operation and it was a pity that it had been laid aside."[9]

Air Engines — Type III — The third class of air engines was the external combustion type operating with an open cycle, some of which employed a regenerator. The man most closely associated with this type was John Ericsson. He was born in Sweden in 1803, came to England in 1826, and finally settled in the United States in 1839. Like Cayley, he was a remarkable inventor. Unlike Stirling, he was a skilled promotor. In this country Ericsson is best remembered by his design of the armored warship, the Monitor, his perfection of the screw propeller, and his "caloric" engines. Apparently Ericsson, as Stirling, believed that the heater surfaces of an air engine were only needed to make up for minor heat losses and, in fact, he so stated in his

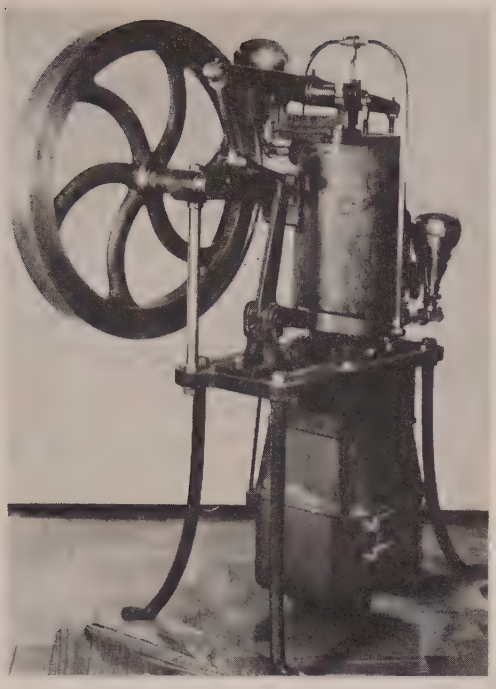


Fig. 23 — Ericsson's Stirling cycle engine (courtesy: Henry Ford Museum)

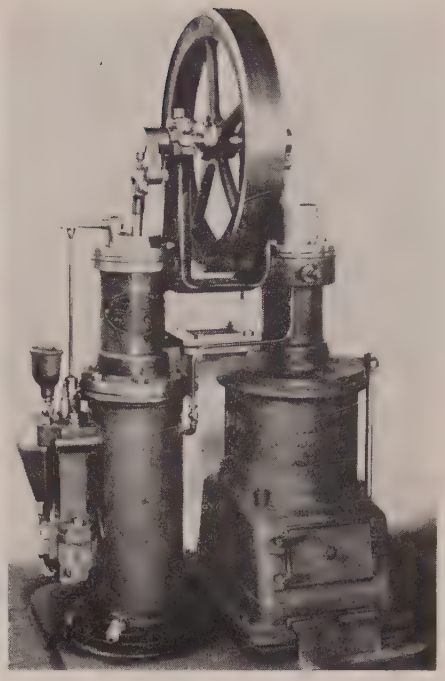


Fig. 24 — Rider engine, 1900 (courtesy: Henry Ford Museum)

patent of 1851.

In 1852 he installed some enormous air engines of his design in an ocean vessel 250 ft long with the expectation that the consumption of fuel would be almost nothing. There were four cylinders, each 14 ft in diameter with a 6-ft stroke, attached to four 11½-ft bore compressing pumps. The engine speed was 9 rpm and the mean effective pressure just over 2 psi. Unfortunately, the ship trials were said to have been conducted in a way which permitted no reliance to be placed on the results. After many revisions and substitutions of parts, the entire enterprise proved to be a failure, and eventually the engines were replaced with steam engines and boilers. Before this, however, interest in the project was world-wide and many discussions of it are found in engineering journals of the day.

Heyday of Air Engine — John Ericsson was by no means ready to give up the hot air engine after the failure of his ship project. He next designed a more or less portable model operating on the open, constant pressure cycle, but without a regenerator. Fig. 22 is from Ericsson's 1859 catalog of the Massachusetts Caloric Engine Co., which he had established about 1857. The opening paragraph of Ericsson's catalog modestly announced the new engine: "This Motor may be confidently pronounced one of the greatest boons which the ingenuity of man has ever bestowed upon his race." It is also mentioned that by 1858 nearly 100 engines had been constructed with cylinders of 8, 12, 18, 24, and 32 in. diameter. During the next few years the engine was sold by the thousands throughout the world and might be considered the first production air engine. Some of the uses to which it was put in the United States included operating machine tools, driving grinding mills, hoisting goods, operating yachts up to 70 ft long, driving printing presses which seemed to be the most common application, and pumping water. Apparently, its advantages were safety, the reduced cost of insurance, and the elimination of the engineer required for steam powerplants.

One of the major disadvantages of Ericsson's caloric engine was its noise. He decided in the 1880's that the silent Stirling cycle offered more possibilities, and his subsequent ventures into the air engine business were for manufacturing and selling his new design operating on the Stirling cycle but without a regenerator. Fig. 23 is a view of one of these engines. In 1898 Ericsson joined with a man from Philadelphia, A. K. Rider, who in 1876, invented an engine which operated on the Stirling cycle. This engine was characterized by its two parallel cylinders with a regenerator in the connecting passage. The Ericsson-Rider engines were sold by the thousands nearly up to the World War I. Being practically noiseless, they were widely used as water pumpers for domestic purposes in homes and buildings. A list of some of the persons who owned air engines reads like a page from *Who's Who* instead of from the Rider catalog: John Jacob Astor, W. K. Vanderbilt, Honorable William E. Dodge, William H. Appleton, Reverend Henry Ward Beecher, J. A. Roosevelt, Governor Leland Stanford. The Rider catalog for 1885 lists water pumpers installed in 113 apartments in New York City, in addition to those installed for 93 domestic and 7 foreign railroad companies, including 21 engines owned by the Michigan Central Railway.

Decline and Fall of Air Engine — The horsepower and efficiency of the Rider engine has been computed from their catalog figures. Their best model, the so-called 10 in. shown in Fig. 24 produced approximately 9/10 bhp with a kerosene consumption of approximately 10 lb per bhp-hr, or about 1.4% brake thermal efficiency. The weight per horsepower was more than 4000 lb while the cubic feet per horsepower was about 120. With this performance it is somewhat amazing that sales prospered as long as they did, but there was litle competition from other silent powerplants; and until electricity

was better distributed and a more central water system installed, these engines did meet a very difinite need. Compared to Stirling's engines, however, they represented a step backwards, for the 1840 Stirling engine of 2 bhp produced over twice the power for one-third the displacement of the Rider engine and was more efficient.

It has not been generally known that multicylinder Stirling cycle engines incorporating double-acting power pistons were successfully constructed after 1850, especially in this country. However, an 1889 issue of the magazine, *Engineering News,* contains a description of such an engine under the title of "A Remarkable New Motor." This engine is shown in Fig. 25. The article did not mention that this 25-hp machine worked on the Stirling cycle, but the authors may not have been aware of this. About the best thing that can be said for the engine was that it utilized all the advantages of the last Stirling engine of 1845 including a regenerator and increased internal pressure. Test run at Boston showed the brake efficiency to be between 8 and 9%. Unfortunately, the designers also made Stirling's mistake of neglecting to include extended heating surfaces.

A few air engines were manufactured in the United States after the World War I. Recently the authors were shown a small Stirling-engine-driven cooling fan, built by a Chicago company until the middle twenties, when it was closed for nonpayment of taxes. Between 1933 and 1939, the Kessler Co. of Chicago manufactured small vertical air engines of the Stirling type with a kerosene heater and an aircooler. The engine appeared to be similar to the Heinrici motor, of which many were built during the late 19th century.

Perhaps one of the most amazing prime movers of all time was the Malone engine, built in England about 1927. It employed all the essential elements of the Stirling cycle with separate cylinders, but instead of a gas it used water which was actually compressed and expanded in the *liquid phase.* Also tried were mercury, oil, and liquid CO_2 and SO_2. The peak cylinder pressure was said to be about 11,000 psi with a mean effective of 2,000 psi. One of the three different working engines constructed appeared to be more than 18 ft high and the indicated efficiency was reported to be 27%.

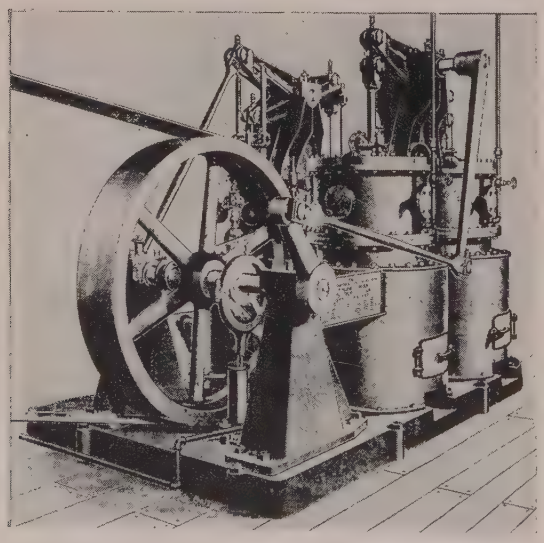


Fig. 25 — "Remarkable new motor" of 1889

Revival of Air Engine — In 1937, the N. V. Philips' Laboratories at Eindhoven, Netherlands, was searching for a portable silent engine having no radio interference for driving small electric generators. The power was needed to operate Philips' radio equipment in remote regions having no electrical supply.

Other than steam plants, which were not very portable, the search narrowed down to the air engine. Several historic engines operating on the Stirling cycle were obtained and tested. The discrepancy between the brake thermal efficiency and the potential Carnot efficiency was so great that Philips felt that the engine warranted some investigation. A research program was started to determine the reasons for the poor results. The studies also included thermodynamic analyses of the cycle and new methods of heater, regenerator, and cooler construction.

During the German occupation the work was considerably restricted. In 1945, one engine was carried off to Germany and discovered by an American military intelligence group soon after the war. The first published work from Philips' was in 1946 by H. Rinia, who is now a director of the Philips Laboratories. Some of Philips' early engines included opposed piston types, V types, some with double-acting pistons, and barrel types with double-acting pistons connected to a swash-plate drive. Designing the piston drive mechanisms often proved to be as difficult as designing compact heat exchangers. Much work was done to improve regenerator effectiveness by the use of very small diameter crimped wires. Some views of the early post war Philips' models and their heat exchanger construction are seen in Philips Technical Review.[10,11] The opposed piston design proved least fruitful, principally due to higher friction because of loading on both pistons. It was also difficult to balance this type in any simple manner. The V designs were easier to balance, and pressurizing the crankcase had raised the mechanical efficiency somewhat. However, the dead volume was excessive. The multicylinder, double-acting design had much to commend it by the elimination of a separate displacer piston, but the pistons had to seal under high-temperature conditions and the dead volume also was greater due to the connecting pipes between cylinders. Eventually the efforts turned to the present single-cylinder, two-piston type in which only one piston is loaded, as in Stirling's 1816 patent.

In 1946, a cooperative program was begun between North American Philips in New York and the United States Navy Bureau of Ships to develop a fractional horsepower single-cylinder Stirling type engine. In 1947, another cooperative program was begun between Werkspoor N. V. of Amsterdam and Philips'. They formed a new combined company under the title N. V. Thermomotor to handle the higher horsepower developments. In 1948, Philips' designed a small single-cylinder engine with an aircooler, kerosene burner and controls, all

[10] "Air Engines," by H. Rinia and F. K. DuPré. *Philips Technical Review,* Vol. 8, 1946, pp. 129–136.

[11] "Philips Air Engine," by F. L. Van Weenan, et al. *Philips Technical Review,* Vol. 9, 1947, pp. 97–104, 125–134.

mounted in a frame with a 200-watt generator. A pre-production run of several hundred units was completed; but as a result of its high cost and the invention of the battery powered transistor radio, the powerplant was never placed on the market. By this time, Stirling engines had been constructed with indicated efficiencies of over 20% and outputs of 15 bhp, but it was felt that much more development work was required.

During this same period, a study by another group of Philips' scientists and engineers was begun on a Stirling refrigerator. Their designs proved to be very successful. Liquid air was first produced by one of these machines in 1950 and a production model was on the market in 1955 for use by laboratories. They are described by Köhler.[12]

With the experience gained during the refrigeration work, a fresh approach to the hot gas engine was taken in 1954. Several major changes were made in the designs compared to those of previous years. Hydrogen was used exclusively in place of air. Because of its lower viscosity and higher heat capacity, it reduced the pumping losses associated with air and the engine efficiency was increased by some 50% or more. In addition hydrogen suppressed the pyrolysis of lubricating oil. With air as the working fluid, the slight amount of oil which might bypass the piston would deposit in the regenerator, obstructing the narrow passages until the engine eventually failed to operate. New techniques for welding and brazing the heat exchangers were also developed, and the heater was removed from the cylinder head and constructed as a series of small tubes between the displacer cylinder head and the regenerator. The early regenerators were of the annular type and some losses due to bypassing as a result of thermal distortion were experienced. The new regenerators were cylindrical and as many as eight to twelve were arranged symmetrically around the cylinder. This also eliminated the high temperature stress problem in the region of the cylinder and regenerator flange. Finally, a radically new drive mechanism incorporating twin crankshafts was invented and this proved to be the answer to the balancing problem.

The combination of these factors led to "breakthrough" for the Stirling cycle engine. By 1958 Philips' had achieved a brake thermal efficiency in a single-cylinder, 40-hp machine, of 39%, and had successfully operated a 4-cyl engine developing 350 bhp. The description of the so-called "rhombic drive" and the 40-hp machine is fully covered by Meijer.[1]

Additional References

1. Patent Specification of 1827, No. 5456, by R. and J. Stirling.
2. Patent Specification of 1840, No. 8652, by J. and R. Stirling.
3. *Encyclopaedia Britannica,* 1957 Edition.
4. Catalog of Caloric Engine Co., 1859. Pub. by W. and E. Howe, Boston.[13]
5. Catalog of Rider Pumping Engine for 1885, Sayer and Co., New York.[13]
6. Catalog of Rider-Ericsson Engine Corp., New York City, for 1900-1910.[13]
7. *Engineering News* (U.S.A.), Sept. 14, 1889, pp. 242-246.[13]
8. "Malone Engine." *The Engineer,* Vol. 152, 1931, pp. 97-101.
9. "Study and Review of Hot Air Engines," by H. H. Bowie. ASTIA Report AD No. 45429, April, 1954.
10. "Philips Air Engine." *The Engineer,* Vol. 184, 1947, pp. 549-550, 572-574.

DISCUSSION

Stirling Engines Shows Promise for Military Applications

— James H. Horton
U. S. Army Engineer
Research and Development Laboratories

FOR THE PAST several years the Army has had an urgent need for small, portable, essentially noiseless powerplants for forward area surveillance and communications centers as well as other special applications. These requirements have been met on an interim basis by the use of conventional internal combustion engines with silencing housings around them. This has not been a really satisfactory solution in that the use of acoustical housings has added considerable weight and bulk to the equipment and the degree of noise attenuation achieved has been far less than desirable. For example, the 3-kw silenced engine generator set meeting current military specifications, although powered by a lightweight commercial aircooled gasoline engine, weighs 333 lb (about 50% more than its unsilenced counterpart), and is audible from 150 yards or less with relatively low background noise. What we are seeking is a smaller and lighter unit which is inaudible at 100 ft.

We have long felt that the closed cycle engine offers one of the most promising approaches to a silent power unit. Our efforts along this line have been directed toward both the vapor cycle engine and the Stirling cycle engine. As long as 12 years ago the Navy tested some Stirling engines of ¼-hp size driving 125-watt generators. Several years later, in 1954 and 1955, similar engines on 180-watt generator sets were tested by the Army Signal Corps. None of the tests were completely successful and the engines were not adopted for military use, primarily because of their excessive size and weight, low efficiency, lack of reliability, and excessive maintenance requirements.

The new Stirling engines as described here appear to represent somewhat of a breakthrough. We are especially impressed with their good thermal efficiency, which is much higher than the earlier models. Another significant advancement is the rhombic crank mechanism which provides almost perfect balance and eliminates piston side thrust. We are also glad to see the increase in speed to 3600 rpm (as compared to 1500–2500 rpm for the earlier models) as this will facilitate development of lighter weight equipment. The possibility of multifuel operation is, of course, another attractive feature from a military standpoint.

On the negative side, the complexity and relatively high cost of the Stirling engine are of some concern, and we hope to see these factors improved with further development. Other problems include the development of effective seals for the working fluid and development of a suitable governing system for generator application. In this con-

[12] "Gas Refrigerating Machine," by J. W. L. Köhler, et al. *Philips Technical Review,* Vol. 16, 1954, pp. 69–78, 105–115.
[13] References courtesy of Henry Ford Museum, Dearborn, Mich.

nection I would be interested in hearing more about the control system developed by Philips.

All things considered, the Stirling cycle engine appears to be one of the most promising approaches to the Army's requirement for silent powerplants.

Multicylinder Stirling Design Presents Weight Problems

— Dr. Simon K. Chen
International Harvester Co.

I AM VERY much interested and excited to see the numerous recent attempts to develop novel engines. Some time ago, we learned that Ford was working on a multicompression, intercooled, regenerative gas turbine engine. Here, we learned about the GMR-Philips hot gas engine, which is a closed cycle, external combustion, regenerative piston engine. These two engines are quite different in design and characteristics. Ford's engine runs at very high rotating speed (around 60,000–70,000 rpm), and it is quite compact. The GMR-Philips engine, on the other hand, is a rather low speed (around 600 fpm piston speed), and it is quite bulky and heavy. However, both engines have one similar objective: they are trying to approach the thermal efficiency of the Stirling-Ericsson engine.

Because of the quietness of the Stirling engine, its potential in good thermal efficiency and its multifuel characteristic, one can be reasonably sure that the Stirling engine will be successful in some applications. To help understand this new engine better, I would like to ask the authors to discuss the following points:

1. Speed and Performance — As this engine is basically an external combustion engine, the power output as well as the performance will be limited by heat transfer and, therefore, by engine speed. What is the analytical relationship between the speed and performance? What is a practical maximum piston speed in fpm?

2. Acceleration and Controls — As an external combustion engine, the acceleration response will no doubt become a problem if only fuel control is used. This slow response might not affect a power unit application, but it will for an automotive or marine application. One way out is to change the pressure level of the hydrogen in the closed cycle. This would require additional piping, pump, and surge tank. Can engine speed and load be independently controlled in this engine? How are the fuel throttle and hydrogen density controls interconnected?

3. Hydrogen as a Medium — To reduce the aerodynamic losses and to enhance the heat transfer, the use of a light inert gas such as hydrogen is necessary. This requires extra tanks and handling. These extra requirements might present some problems if the engine is used as a rural powerplant as Philips first proposed.

4. Multiple Cylinders — Due to the rhombic drive, external burners, and the rest, very little weight or space can be saved if a multicylinder Stirling engine is made. A multicylinder Stirling engine must be made by stacking single cylinders together. In other words, the usual weight saving on the lb/hp basis in going from a single cylinder to a conventional multicylinder engine, does not exist with the present GMR-Philips design.

Authors' Closure To Discussion

IN REPLY to Mr. Horton, the control mechanism consists of a pressure control which almost instantaneously bleeds the pressure of the working fluid to atmospheric pressure for a sudden reduction in load. A separate fuel control is actuated by a temperature sensing element to maintain constant cylinder head temperature. Increases in load are taken care of by pumping the working fluid back into the cycle thus raising its mean pressure.

In response to Dr. Chen I would point out that our engine *is* a Stirling cycle engine and the high efficiencies which have been achieved have only been achieved with the Stirling cycle not the Ericsson engine. With regard to engine weight they are already quite competitive with high output diesel engines (in the neighborhood of 10 lb/hp). There is no restricting relationship between speed and performance, since the design can be chosen to peak at any reasonable crankshaft speed and because of the short stroke there appears to be no practical limitation in piston speed. For example, we have run a small model of the engine as high as 5000 rpm.

Dr. Chen is quite right in mentioning the problem of acceleration response if only a fuel control is used. However, the pressure control as outlined above is quite satisfactory.

In order to increase the pressure level of the working fluid an accumulator or storage tank and separate pump are required. The size of this auxiliary equipment is dictated by the response characteristics desired. If rapid increases in load are anticipated a larger displacement pump and accumulator would be used. There is no direct interconnection between the fuel control and the pressure control since the fuel is modulated by a temperature control as outlined above.

The use of hydrogen does, of course, reduce the aerodynamic losses to a minimum, however, it also eliminates the oxidation of the lubricating oil which would form deposits on the regenerator thus reducing its efficiency.

With regard to multiple cylinders I see no reason why the same weight saving that occurs in other engines cannot be achieved with the Stirling engine since a reduction in size and optimization of engine driven accessories can always be made as is commonly done with internal combustion engines. The largest Stirling cycle engine which is now running is 380 hp and with no concerted effort to reduce weight is approximately 10 lb/hp.

ORAL DISCUSSION

Reported by O. A. Uyehara
The University Wisconsin

S. I. Macduff, Bendix Aviation Corp.: In a conventional engine with the crankshaft on the center line of the cylinder, the diagram of piston motion versus time or crank angle is symmetrical, although not purely sinusoidal. A similar diagram for the engine with a rhombic drive, which in effect would merely provide for the crank centerline being displaced away from the center of the cylinder, will be highly asymmetrical.

Was the rhombic drive adopted for the purpose of making use of the asymmetry? If so, how much advantage is gained thereby, or alternatively if the drive is adopted solely for the stated purpose of eliminating side loadings on pistons?

Mr. Flynn: Mr. MacDuff is quite correct that the piston motion is no longer the same as for an engine with piston operating on the centerline of the crankshaft. However, it is necessary to get phase relationship between the displacer piston and the working piston in order for the cycle to work and the rhombic drive achieves this phasing with inherent perfect balancing.

John M. Bailey, Caterpillar Tractor Co.: What means is used to lubricate the power piston in the hot gas engine and has any trouble been encountered with the sealing ring?

Mr. Flynn: The power piston is lubricated by a force feed lubricator similar to that used in large diesel engines. No particular trouble has been encountered with the seals; however, these are an area for continued development since any leakage by the seals results in a loss of the working fluid which must be replaced. One of our concerted research efforts is directed to improving seal materials and designs.

DEVELOPING TRANSAXLE FLUID

EXTENSIVE TESTING by GM Research Laboratories has screened five promising transaxle fluids out of 32 mineral-oil-base fluids, 10 synthetic-base fluids, and numerous additive-base stock combination fluids. This paper discusses the findings of the testing and the continuing program on the five fluids.

Transaxle fluids have a number of properties affecting performance, including:

High-temperature viscosity.
Low-temperature fluidity.
Shear resistance.
Friction properties.
Oxidation resistance.
Antifoam quality.
Effect on seals.
Fluid-clutch plate compatibility.
Antiwear quality.
Extreme-pressure quality.
Antirust and anticorrosion qualities.*

N. A. Hunstad, R. A. Wilkins, R. E. Osborne, and E. D. Davison

Research Laboratories, General Motors Corp.

IN COMBINING the transmission and axle mechanisms into a transaxle, it is logical to consider the use of a single lubricant rather than one for each of the two parts of the unit. There are many advantages in favor of the single fluid, including:

1. Rotating-shaft seals between the transmission and axle portions would not be required. The result would be greater design freedom, cost reduction, and absence of production and service problems associated with the seals.

2. The use of only one fluid would simplify car production and fluid servicing.

The problems of developing such a single lubricant depend on the nature of the transaxle. Its requirements will differ somewhat, for instance, depending on whether spiral bevel or hypoid gears are used. Even greater differences would be involved between manually shifted and automatic types. Present indications are that both types will be in use and that both will employ hypoid gears.[1]

While the lubricant for the manually shifted type should present no problem with the present knowledge of gear lubricant, there are many problems in the development of a fluid for the automatic type. That these problems should arise is not surprising to one who knows the history of hypoid gear lubricant and automatic tranmission fluid. It is the purpose of this paper to review this history and to discuss the efforts of GM Research Laboratories in the development of transaxle fluid for the automatic-type unit.

History of Hypoid Gear Lubricant

Development of the Hypoid Gear — One of the earliest references to hypoid-type gears was made in *Machinery* in 1923.[2] The term "spiral hyperboloidal gear" is used in this reference. The concept of hyperboloidal pitch surfaces was well described by Griffith in the General Motors Engineering Journal.[3] According to this reference, "The shorter term (hypoid) is merely a contraction (of hyperboloidal) coined many years ago and properly belongs to the Gleason Works which makes the machine tools for commercial manufacture of this type gear."

The advantages of the hypoid gear over the previously used spiral bevel gear have long been recognized. At least as early as 1926 gear engineers[4] said, "The chief advantages of hypoid gears are noiseless operation, increased load-carrying capacity, the possibility of high reduction and low numbers of teeth, long life, and high efficiency." Also, while many consider the trend toward lower car bodies a relatively recent development, this design potential in the lowered pinion position was recognized by these same engineers: "The use of a ring

* Paper presented at SAE Annual Meeting, Detroit, Jan. 11, 1960.

1 "Development of Transaxle Fluid for Chevrolet Corvair," by R. E. Harvie, E. L. Nash, and J. W. Clark. Paper presented at SAE Annual Meeting, Detroit, January, 1960.

2 "Hobbing Spiral Bevel Gears," by Nikola Trbojevich. *Machinery*, vol. 23, Nov. 22, 1923, pp. 225–235.

3 "Gear Wheels and Applied Mathematics," by Bain Griffith. *GM Engineering Journal*, Vol. 1, Sept.–Oct., 1953, pp. 32–39.

4 "Design, Production and Application of Hypoid Rear-Axle Gear," by A. L. Stewart and Ernest Wildhaber. SAE Transactions, Vol. 21, Part II, 1926, pp. 386–411.

gear of smaller diameter and the location of the rear end of the propeller shaft will have an important influence on chassis design. On account of the capacity of hypoid gears for greater transmitted load, a decrease of about 10% in the ring-gear diameter can be made without increasing the unit stresses. This makes possible a greater road clearance. The lowered position of the driving pinion removes the chief obstacle to lowered floorboards at the rear end. Advantage can be taken of this to lower the entire body."

Early Packard Experience — Hypoid gears were first used in production passenger-car rear axles in the 1926 model Packard.[5] The service lubricant recommended[6] by Packard at that time was "a good real axle fluid mineral gear oil." It was further recommended that "in cold weather the lubricant should be thinned with cylinder oil to approximate its summer consistency." These recommendations were the same as those for the spiral bevel gears used previously by Packard.

The use of mineral gear oil is rather surprising since it is now known that extreme-pressure additives are required in hypoid gear lubricants. Apparently, Packard soon learned of this requirement as judged by their early experiences in the lubrication of hypoid gears:[7] "Before adopting them we made careful analytical comparisons of tooth pressures, sliding factors, surface stresses, and a hundred other items; but we passed over the lubricant entirely. But for a fortunate circumstance prevailing at the time, this whole achievement would not have been realized. It happened that about a year previous to all this, we had been induced, after careful testing, to adopt a new rear-axle lubricant about which all technical details were held secret by the maker. It smelled like creosote, and was black and viscous; and it was guaranteed to make gears run quietly, 'stand up' better, reduce backlash, and permit the gear experts to sleep nights. As a matter of fact, it was supposed to do everything but cure snake bites or relieve lumbago. It actually happened to be an 'EP' (Extreme Pressure) lubricant of very high EP value, and was later found to be a lead-soap corrosive sulfur compound. Little did we suspect the high virtues of this oil, and we went serenely on and used it in these hypoid axles. There is not the slightest doubt in my mind about what would have happened had we use a straight mineral oil on these tests, and we can only speculate on what the present would hold for hypoid gears. But we were not long in discovering the need for this lubricant. Our first experience was to be had when some of our less cautious owners were sold on an oil change for their rear axles at the corner greasing station instead of at the Packard service station. We learned then about scored gears, and the following winter taught us some lessons on pour-test requirements and lead-soap deposition." It is of interest that in the 1929 Packard owner's manual the rear axle lubricant recommendation was changed to read: "Use Whitmore's 'O'." It is understood that this product was a lead-soap active-sulfur gear oil.

"Extreme-Pressure Lubricants" — Packard's experience with hypoid gears was paralleled by that of engineers in other groups, so that there was soon a general recognition of the special antiscore requirement of these gears. According to Mougey,[8] ". . . the Lubricants Committee of the Society of Automotive Engineers in 1931 gave to the lubricants that are required for these more severe conditions (high sliding velocity and high load) the name 'Extreme Pressure Lubricants'."

Evaluation of EP Quality — There have been many efforts to develop an EP test. A four-square type machine using two standard production axles was employed by Gleason at least as early as 1926. According to Witham,[5] this machine was used specifically for the testing of hypoid lubricants for load-carrying capacity. In 1926 Cadillac[9] set up a four-square axle test in which four production axles were used. A similar test was described by Wooler[10] in 1929. He also discussed a "laboratory roadtest" in which an engine-driven axle was loaded with flywheels.

Griswold[7] described a road test for evaluating lubricants: "A common and extremely severe test for scoring is the following: The car is driven to a speed of approximately 60 mph and allowed to coast, while the transmission is shifted into second gear; then, with the engine idling, the clutch is suddenly re-engaged as the speed reaches 50 mph. This kind of test is quick and convenient to make, and is a good criterion for determining the relative merits of a number of lubricants as to EP value in a very short time." This type of test is often referred to as a "bump test." Through the years many similar tests have been developed. For instance, in the General Motors Manufacturing Standards numbers GM 4743-M and GM 4744-M five such tests are listed:

1. Buick Motor — Procedure 10-A.
2. Cadillac Motor — Procedure 4-A.
3. Chevrolet Motor — Procedure 425-B.
4. Oldsmobile — Schedule 26.
5. Pontiac Motor — Engineering Dept. Manual, Test 25.

No doubt there have been many other full-scale axle tests used through the years. However, being aware that full-scale gear tests were time-consuming and costly, many investigators used bench tests in their efforts to evaluate gear lubricants. For instance, in 1931 The Timken Roller Bearing Co. anannounced an apparatus[11] for testing the load-carrying capacity of lubricants. In the same year, Mougey and Almen[12] reported the results obtained with a bench test which they used to investigate ". . . the problem of preventing galling or scoring between heavily loaded sliding surfaces." They were able to demonstrate the effectiveness of EP additives with their machine. In 1932, Wolf and Mougey[13] reported on the correlation of the Almen Extreme-Pressure Lubricant Testing Machine data with service performance. While they obtained

---

[5] "Hypoid Gears, Axles, and Lubricants," by W. A. Witham. SAE Transactions, Vol. 32, 1937, pp. 509–513.

[6] "Information Packard Motor Cars" (owner's manual), Packard Motor Car Co., 1926, p. 19.

[7] "Rear-Axle Design and Lubrication," by W. R. Griswold. Paper presented at API meeting, May, 1936.

[8] "Extreme Pressure Lubricants," by H. C. Mougey. *Science of Petroleum*, Vol. 4, 1938, pp. 2598–3001.

[9] "Dynamometer Testing of Gears in Normal Environment," by G. L. Rothrock. Paper presented at SAE National Tractor Meeting, Milwaukee, September, 1956.

[10] "Laboratory Testing of Rear Axles," by E. Wooler. SAE Transactions, Vol. 24, 1929, pp. 346–355.

[11] "Apparatus Tests Load Capacity of Lubricant," *Refiner and Natural Gasoline Manufacturer*, Vol. 10, 1931, p. 140.

[12] "Extreme Pressure Lubricants," by H. C. Mougey and J. O. Almen. Paper presented at API meeting, November, 1931.

[13] "Extreme-Pressure Lubricants," by H. R. Wolf and H. C. Mougey. Paper presented at API meeting, November, 1932.

good correlation with some lubricants, they stated that: "Certain lubricants which will not carry the full load on the Almen machine may function satisfactorily under service conditions, especially if the service is not too severe. However, of the lubricants which do not carry the full load on the Almen machine we do not believe that the ones which fail at moderate loads are of necessity inferior to the ones which fail at somewhat higher loads. We are now working on modifications of the method of operating the machine to try to obtain better correlation with service in this respect."

McKee, Harrington, and McKee[14] evaluated nine lubricants in four bench test machines reported[15] to be: (1) Almen machine, (2) Timken machine, (3) Floyd machine, and (4) G.M.R. Lubricant Testing Machine. According to McKee, Bitner, and McKee,[16] in the former work ". . . no set of operating conditions was found for each machine such that all the machines would rate all the lubricants in the same order, . . . and no one machine rated the lubricants in the order of their service performances." With the machine developed at the Bureau of Standards, McKee, Bitner, and McKee reported, ". . . the test results obtained with this machine appear to rate the lubricants in reasonable agreement with ratings based upon service performance in automotive gears." Unfortunately, they were probably somewhat over-optimistic, since the SAE machine developed by them with modifications by James and Smith[15] is not generally relied upon today for gear lubricant evaluation.

As indicated in the literature there have been many frustrating efforts to develop a meaningful EP bench test. A story which illustrates what is probably the fundamental reason for the failure of these efforts was given by Almen:[17] "Trouble is that test machines like animals have different tastes. A horse prefers hay, a dog likes meat, and a squirrel likes nuts. If we want to compare hay, meat, and nuts using a horse as the testing device, there will be no doubt as to which food is best. But if a neighbor makes the same comparison basing his opinion on the reactions of a dog, he will find that hay and nuts are inedible and only meat can be rated as a food. To settle the resulting controversy an arbiter is chosen. This man will not employ the services of a horse or dog, because they have been found unreliable, so he tries the three foods on a squirrel and the decision is so clear that there is no longer any doubt as to the best food. Nuts!" Just as horses, dogs, and squirrels have different tastes, so apparently do EP test machines and axles. Consequently, it is not surprising that the various test machines and axles rate lubricants differently.

However, a low-cost EP evaluation technique is still a very desirable objective and there will no doubt be further efforts to achieve it. An example of such an effort was presented in 1959 by DeArdo and Sargent.[18] They discussed their procedure for evaluating heavy-duty gear lubricants by using three bench-test devices and combining the three results to establish ratings for lubricants. They show that their procedure classifies lubricants according to load-carrying capacity. However, they do not show any correlation between their procedure and full-scale axle tests.

The overall impression gained in reviewing all of these efforts to measure EP quality in full-scale axle tests and in bench tests is that the full-scale tests are the only reliable ones. As expressed by Zwahl,[19] "In the automotive industry it is our belief that the only suitable test by which we can ascertain whether or not a gear lubricant will satisfactorily lubricate a passenger-car hypoid axle at high speeds is to put the lubricant in the axle under the car and run the car at high speeds."

Large-Scale Use of Hypoid Gears — According to Mougey,[20] until 1936, "The number of cars with hypoid axles was too small to justify stocking the lubricants (hypoid gear) by the average filling station." This situation changed when three General Motors divisions began using hypoid gears in the 1937 models and the remaining two adopted this type of gear in their 1939 models.

With the introduction of the 1937 models, General Motors began supplying hypoid lubricants through the car dealers. At the same time they began developing an approved list of lubricants.[21] Satisfactory lubricants soon became generally available and the list was discontinued, being replaced by a specification dated July, 1937.

The field lubricants were generally of the lead-soap active-sulfur type. While they provided satisfactory passenger-car performance, it was found that they gave excessive wear in heavy-duty truck hypoid gears.[22] Consequently, it was necessary to have a different type of lubricant to meet heavy-duty truck requirements. However, some lubricants were available which were claimed to function satisfactorily in both passenger cars and trucks.

Military Gear Lubricant Specifications — The 1940 decision of the Army[23] ". . . to utilize one gear oil for all types of automotive gear drives . . ." was a natural development. In 1942, Federal Specification VV-L-761 for Universal Gear Lubricant (U. S. Army Specification 2-105A) was issued in proposed form.[24] It included a durability test using a truck axle and a shock test conducted in a car on the road. It also included an SAE Machine test. This specification required ". . . a load-carrying additive containing sulphur and chlorine . . ." Due to its restrictive nature and the tendency of VV-L-761 lubricants to sludge at high temperature, the VV-L-761 specification was replaced by U.S. Army Specification 2-105B in 1946. It is of interest that the latter included a high-speed car test and a high torque axle dynamometer test, but no SAE Machine test or elemental chemical composition requirement.

With only minor modifications the 2-105B speci-

---

[14] "Load Carrying Capacity of Extreme-Pressure Lubricants," by S. A. McKee, E. A. Harrington, and T. R. McKee. SAE Transactions, Vol. 28, 1933, pp. 217–223.

[15] "Extreme-Pressure Lubricants Testing," by G. L. Neeley. SAE Transactions, Vol. 31, 1936, pp. 293–296.

[16] "Apparatus for Determining Load-Carrying Capacity of Extreme-Pressure Lubricants," by S. A. McKee, F. G. Bitner, and T. R. McKee. SAE Transactions, Vol. 28, 1933, pp. 402–408.

[17] General Motors company correspondence, J. O. Almen to D. C. Wilkerson, Sept. 6, 1949.

[18] "Laboratory Rating Method for E. P. Lubricants," by A. J. DeArdo and L. B. Sargent, Jr. Paper presented at First American Society of Lubrication Engineers Gear Symposium, January, 1959.

[19] "High-Speed Testing of Gear Lubricants for Hypoid Axles," by C. E. Zwahl. Paper presented at SAE Summer Meeting, French Lick, Ind., June, 1946.

[20] "Hypoid Gear Lubrication," by H. C. Mougey. Paper presented at National Lubricating Grease Institute, October, 1937.

[21] "Hypoid Lubricants — Test Results and Their Interpretation," by C. E. Zwahl. SAE Transactions, Vol. 32, 1937, pp. 555–565.

[22] "Hypoid-Lubricant Requirements for Passenger-Car and Truck Rear Axles," by H. R. Wolf. API Proceedings, Vol. 20M-20, Section III, 1939, pp. 16–24.

[23] "Recent Developments in Universal Gear Lubricants," by P. V. Keyser. Paper presented at SAE Summer Meeting, French Lick, Ind., June, 1946.

[24] "Recent Developments in Automotive Gear Lubricants," by W. B. Bassett. Paper presented at SAE Chicago Section meeting, May, 1947.

fication was subsequently identified as MIL-L-2105.[25] Lubricants qualified under this specification gave satisfactory performance in military vehicles for some time.[26] However, field difficulties occurred which led to conducting of tests in military trucks at the Yuma Test Station in 1953. Lubricants meeting the MIL-L-2105 specification failed due to deposit formation and ridging and scoring of gear teeth. These observations were confirmed in tests conducted in subsequent years. These results led to considerable activity on the part of additive manufacturers in the development of new gear oil additives.[27, 28, 29, 30]

Also, during this time the automobile manufacturers urged a review of gear lubricant qualities for high-speed operation. This activity was aided materially by the development by Sands[31] of a reference oil system with which axle gear antiscore requirements could be measured. For instance, in 1956 General Motors ran tests on passenger cars and established that for new gears 14–90 performance level is required and for broken-in gears 10–90 is required. It should be pointed out that these requirements were established using noncoated gears.

Using the information developed by Ordnance and by the car builders, the CRC developed new axle dynamometer test techniques: L-37 to simulate heavy-duty gear operation, and L-42 to simulate passenger-car operation. These two tests were used as the basis for the present MIL-L-002105A (MIL-L-2105B) specification introduced in 1959. Suppliers are now qualifying oils against this specification.

Multipurpose Type Gear Lubricant (API Service GL 4) — During the development of the new type gear lubricants, the American Petroleum Institute promoted the development of the name "Multipurpose Type Gear Lubricant (API Service GL 4)" and the associated definition.[32, 33] The definition describes the type of service for which GL 4 is intended; no performance tests are included. Consequently, it is subject to considerable variation in interpretation. It is understood, for instance, that lubricants of about the 6 performance level are being labeled for GL 4 service, whereas it was the understanding of many that 10 performance level was to be the minimum quality for lubricants for API Service GL 4.

Furthermore, there have been reservations about the performance even of products compounded to the 10 performance number level. For instance, at the National Petroleum Association Meeting in 1957, several speakers[34, 35, 36] voiced apprehension about the performance of the new gear lubricants. Also, several heavy-duty equipment manufacturers have urged caution in the use of the new lubricants or even recommended against them for some types of service.

However, some of these reservations have been dispelled, as indicated by the fact that four out of five of the General Motors car divisions now use a lubricant formulated with the new type additive for either part or all of their axle initial fill. Also, Johnson[37] indicated favorable results with the new lubricants in recent experience with heavy-duty vehicles.

Despite the considerable favorable experience with GL 4 lubricants, there are still many automotive manufacturers who do not call for them in their owner's manuals. At least part of the reason for the refusal of some to recommend them is the wide range of quality of GL 4 lubricants in the field. The remedy for this situation is being sought by an ASTM Subcommittee. Perhaps the performance tests used by the equipment manufacturers can be used as a basis for a sound GL 4 lubricant definition.

Hypoid Gear Lubricant Qualities Required in Transaxle Fluid — The literature shows that most of the hypoid gear lubricant development has been directed toward antiscore quality. Other factors such as bearing wear, rust, corrosion, viscosity, heat resistance, seal compatibility, and the like, must also be considered in evaluating gear lubricant. Problems of matching lubricant qualities with the requirements of hypoid gears still exist and are receiving much attention.

All of these gear lubricant qualities must be considered in the development of transaxle fluid. This is a logical conclusion since the transaxle will employ hypoid gears in a service similar to that in which current rear axles are used. In addition, for the automatic-type unit, transaxle fluid must also have the qualities of automatic transmission fluid which have been developed during the past 23 years.

History of Automatic Transmission Fluid[38]

Current automatic transmission fluid has evolved as experience has been gained with automatic transmissions in the field. For instance, when engine oil was used in the "automatic safety transmission" in 1937 it was found that improvement in the oxidation resistance of the lubricant was required. Then, as engineers worked with better fluids they became aware of the need for careful control of frictional characteristics to provide satisfactory clutch engagement and durability. Another requirement that became evident was that of antifoam quality.

From these and other early experiences, the transmission engineers came to fully appreciate the need for performance-tested lubricants. Accordingly, automatic transmission fluids were supplied through car dealers only, up until 1949. In obtaining a wider distribution of service fluids it was necessary to make certain that fluids supplied through other outlets were also performance tested. The engi-

[25] "Development of Automotive Gear Lubricants — Present and Future," by T. P. Sands and J. L. Werner. Paper presented at American Society of Lubrication Engineers meeting, September, 1954.
[26] "Performance Testing of Gear Lubricants in Military Equipment," by R. E. Streets. Paper presented at SAE National Fuels and Lubricants Meeting, Tulsa, November, 1956.
[27] "Chemical Additives for Hypoid Gears," by C. V. Smalheer. Paper presented at American Society of Lubrication Engineers meeting, April, 1957.
[28] Discussion by T. P. Sands. National Petroleum Assoc. Symposium on Gear Oil Developments, September, 1957.
[29] "Some Additional Characteristics of New Multipurpose Gear Lubricants," by W. B. Bassett. Paper presented at National Petroleum Assoc. Symposium on Gear Oil Developments, September, 1957.
[30] "Development and Application of Broad Performance Range Gear Lubricant Additive," by R. K. Williams, W. C. Brandow, and J. W. Schulte. Paper presented at Independent Oil Compounders Assoc. meeting, September, 1957.
[31] "Development of New Automotive Gear Lubricants," by T. P. Sands. Paper presented at American Society of Lubrication Engineers meeting, April, 1957.
[32] SAE Handbook, 1959, pp. 284–285.
[33] "Status of Automotive Rear Axle Lubricants," by C. R. Noll. Paper presented at SAE Metropolitan Section meeting, May, 1958.
[34] "Some Thoughts on Automotive Gear Lubrication," by W. A. Johnson. Paper presented at National Petroleum Assoc. meeting, September, 1957.
[35] "General Motors Looks at Gear Lubricants — September 1957," by N. A. Hunstad. Paper presented at National Petroleum Assoc. meeting, September, 1957.
[36] "Problems Related to Use of High Activity Gear Lubricants," by F. X. Sieloff and G. Hoffmann. Paper presented at National Petroleum Assoc. meeting, September, 1957.
[37] "Lubricating Heavy-Duty Commercial Drive Axles," by W. A. Johnson. Paper presented at SAE Mid-Continent Section meeting, April, 1959.
[38] "Present Status of Automatic Transmission Fluid, Type A," by N. A. Hunstad, T. W. Selby, and R. E. Osborne. Paper presented at SAE Summer Meeting, Atlantic City, June, 1956.

neers, therefore, included full-scale transmission tests in the 1949 target specification for Automatic Transmission Fluid, Type A. A viscosity index of 150 minimum was also included in the hope of obtaining products better meeting transmission needs at temperature extremes. However, the shear stability of fluids meeting the target specification was so poor as to make them impractical. Therefore, the specification was modified in 1950 to cover more shear-stable products.

Field experience in the winters of 1949–50 and 1950–51 showed the need for better low-temperature properties to prevent clutch-plate burning. Thus, the specification was again modified to include a full-scale cold-room test.

Fluids meeting the 1951 specification proved to be satisfactory for a number of years until the controlled-coupling type transmission introduced a more severe oxidation resistance requirement. As a result, the "Suffix A" program was introduced in 1957. Under it, all qualified fluids were subjected to a more stringent oxidation test. Practically all fluids in the field are now of Suffix A quality.[39]

Through the years, petroleum technologists have come to look upon automatic transmission fluid as a necessarily complex product requiring very careful formulation. Bozzelli[40] stated: "The modern automatic transmisson is one of the most complicated mechanisms in today's passenger car. Therefore, it is not surprising that the fluids which are used in these transmissions are similarly complex. It is our belief that automatic transmission fluid is one of the most detailed and precise products made by the petroleum refiner."

Texaco, Inc. discussed some of the functions of automatic transmission fluid:[41] "Even the most casual understanding of the passenger-car Hydra-Matic should include the realization that the fluid used in it must be regarded as a highly essential special part that must perform all of the following important and often diverse functions, and all with very high ability:

1. A mild extreme pressure gear lubricant to protect heavily loaded helical and spiral gearing.
2. A torque fluid with nonfoaming characteristics for fluid coupling operation.
3. A hydraulic fluid capable of permitting the Hydra-Matic to operate between minus 30 and plus 300 F.
4. A wet clutch and transmission band lubricant that facilitates the smooth and noiseless initial engagement of these mechanisms without permitting subsequent slippage.
5. A lubricant with such high oxidation resistance that it can withstand exposure to atmospheric oxygen for 25,000 miles at high temperatures and abundant catalysis without forming sludge or varnish that would prevent proper operation of the Hydra-Matic's many control valves.
6. A rust preventive and anticorrosion lubricant which will not adversely affect any of the many metals, rubber seals, gaskets, adhesives, clutch facings and band liners used in the Hydra-Matic."

These functions are also generally applicable to automatic transmissions other than the Hydra-Matic.

As well as recognizing the complexity of automatic transmission fluid, the technologists in the petroleum and automotive industries have directed considerable effort to the formulation and evaluation of this complex product. At a round table meeting[42] on "Automatic Transmission Fluid Trends" in 1957, the field service picture and laboratory tests were discussed. Foehr and Calish[43] discussed their development of an oxidation bench test and its correlation with the transmission oxidation tests in a 1959 meeting. At the same meeting, Kite and Koenig[44] gave a comprehensive report on transmission fluid requirements and tests used in evaluating fluids. Also, Bozzelli[40] outlined the design targets and considerations in developing a fluid. At a later meeting in 1959, Selby[45] reviewed the development of a low-temperature viscosity bench test and its correlation with a full-scale cold-room test using a car. Crosthwait and Greenawalt[46] discussed the effects of base stocks and additives on low-temperature fluidity. Deen and Stendahl[47] reported on a bench oxidation test employing a filter-paper rating system and on the correlation of the test with transmission oxidation tests. Moreover, they clearly illustrated the effect of specific gravity of rubber compounds on the volume change observed in testing fluids. They also showed the relationship of viscosity loss obtained in a chassis dynamometer test to that in the 50-hr Hydra-Matic Durability Test included in the specification for Automatic Transmission Fluid, Type A, Suffix A Identification.[48]

A point of interest regarding fluid evaluation is that better correlation apparently has been obtained between bench and full-scale automatic transmission fluid tests than has been possible between bench and full-scale EP tests for hypoid gear lubricant. For instance, both Foehr and Calish,[43] and Deen and Stendahl[47] showed correlation in the case of oxidation tests, and Selby[45] demonstrated that low-temperature viscosity can be measured satisfactorily in a bench test. Perhaps the reason for the apparently better success in the case of automatic transmission fluid is that the factors being tested are not so critical of test conditions as EP quality is. Or, to put it another way, in evaluating a product such as drinking water, for instance, the horse, dog, and squirrel referred to earlier might well give a sample equal ratings.

The literature referred to shows that much work has been done to achieve satisfactory fluid performance. While it is true that current fluids are performing reasonably well in the field,[49] consideration is being given to further improvements. In 1959 General Motors issued a proposed specification[50] in

---

[39] GMC list of "Qualified 'Type A' Automatic Transmission Fluids," fifth Edition, July 1, 1959.

[40] "Designing Automatic Transmission Fluid," by A. J. Bozzelli. Paper presented at SAE Annual Meeting, Detroit, January, 1959.

[41] "Hydra-Matic, 1939 Thru 1956," *Lubrication*, Vol. 42, February, 1956.

[42] "Compatible Automatic Transmission Fluids," secretary's report by E. F. Koenig. SAE *Journal*, Vol. 65, October, 1957, pp. 78–79.

[43] "Correlation of Bench Oxidation Tests with Transmission Test Performance of Automatic Transmission Fluids," by E. G. Foehr and S. R. Calish. Paper presented at SAE Annual Meeting, Detroit, January, 1959.

[44] "Evaluating Automatic Transmission Fluids," by W. H. Kite, Jr., and E. F. Koenig. Paper presented at SAE Annual Meeting, Detroit, January, 1959.

[45] "Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance," by T. W. Selby. SAE Transactions, Vol. 68, 1960, pp. 457–467.

[46] "Automatic Transmission Fluids Low Temperature Viscosity Studies," by R. E. Crosthwait and W. F. Greenawalt. Paper presented at SAE National Fuels and Lubricants Meeting, Chicago, October, 1959.

[47] "Oxidation Stability, Shear Stability and Rubber Swell Properties of Automatic Transmission Fluids," by H. E. Deen and C. M. Stendahl. SAE Transactions, Vol. 68, 1960, pp. 448–456.

[48] "Automatic Transmission Fluid, Type A, Suffix A Identification Specification," November, 1957, and two modifying letters dated Aug. 11, 1958 and April 10, 1959. General Motors Corp.

[49] "Type A and Type A, Suffix A, Automatic Transmission Fluids," by W. H. Kite, Jr. Paper presented at API meeting, February, 1959.

[50] Letter on "Proposed Automatic Transmission Fluid Specification," by N. A. Hunstad, to petroleum industry, January 14, 1959.

which improvements in many fluid qualities were outlined. Whether the changes included are economically justified remains to be seen. However, the fact that such a specification was issued indicates that transmission engineers believe that the optimum in automatic transmission fluid performance has not been reached.

Transaxle Fluid

Nature of the Problem — The histories of both hypoid gear lubricant and automatic transmission fluid show that considerable effort has been made to achieve the performance qualities of present products. Even so, they still exhibit performance weaknesses in service. Despite these facts, it is now proposed that a single lubricant be developed incorporating the desirable properties of both hypoid gear lubricant and automatic transmission fluid.

General Requirements — In evaluating automatic-type transaxle fluids, it has been assumed that transaxles will have all of the lubricant requirements of both automatic transmissions and rear axles. The importance of these requirements relative to the fluids normally used in transmissions and axles is presented in Table 1 together with the resultant importance of those for the transaxle.

This table shows that the requirements of the transaxle are similar to those of the automatic transmission. The important exception is that transaxle fluid must have a higher level of EP quality than that required in automatic transmission fluid. This fact is illustrated in Fig. 1. It was recognized that the formulation of a transaxle fluid would be a very difficult problem and that any mechanical variation that would reduce the transaxle lubrication requirement severity would make the solution easier. One possible variation was the phosphate coating of the hypoid gears. It has generally been found that this coating reduced the EP requirement of the gears.[51] Gear scoring tests run on proposed transaxle fluids illustrate this effect. Results obtained in the GMR Hypoid Gear Test for Transaxle Fluid, No. 2 (hereafter referred to as the GMR No. 2 Gear Test[52]) are shown in Table 2.

With the two fluids tested the beneficial effect of the phosphate coating was very pronounced. In the two tests run with noncoated gears, the tests were discontinued due to excessive noise which occurred before five full-throttle accelerations were completed. In comparison, with the phosphate-coated gears, 40 accelerations were successfully completed.

On the basis of such experiences, the decision was made to use only phosphate-coated gears in the work reported here. It is recognized that there is a hazard in such a decision in that certain lubricants might function better with noncoated gears. For instance, in 1956, Southwest Research Institute ran L-37 type tests on coated and noncoated gears using Reference Gear Oil-58-55. Satisfactory results were obtained with the noncoated gears and very unsatisfactory results with the coated gears. Also, in a GMR No. 2 Gear Test conducted by GMC on a commercial SAE 90 grade gear lubricant known to be satisfactory with noncoated gears, unsatisfactory results were obtained using phosphate-coated gears. Thus, while the preponderance of experience indicates that phosphate coating is beneficial, it is possible that a specific lubricant may work better with noncoated gears.

Table 1 — Lubricant Requirements

	Automatic Transmission	Rear Axle	Transaxle
High-Temperature Viscosity	Major	Intermediate	Major
Low-Temperature Fluidity	Major	Intermediate	Major
Shear Resistance	Major	Minor	Major
Friction Properties	Major	Intermediate	Major
Oxidation Resistance	Major	Intermediate	Major
Antifoam Quality	Major	Intermediate	Major
Effect on Seals	Major	Major	Major
Fluid — Clutch Plate Compatibility	Major	Minor	Major
Antiwear Quality	Major	Major	Major
Antirust and Anticorrosion Quality	Intermediate	Intermediate	Intermediate
Extreme-Pressure Quality	Intermediate	Major	Major

Early Experience — The first transaxle fluid evaluated was a mineral-oil-base product obtained from an oil supplier. It was similar to automatic transmission fluid in most respects except that it had a higher level of EP quality.

The fluid was first tested in an axle and was found to give good performance. However, when tested in a Powerglide transmission in a car, malfunction occurred in less than 50 miles of ordinary driving. The cause was severe flaking and glazing of the clutch plates. Subsequent testing of various base stock-additive combinations proved that the EP agent and the clutch plates were incompatible. The adverse effect of the EP agent on the clutch plates is illustrated in Fig. 2.

This and other early experience indicated that hypoid gear protection and clutch-plate lubrication are two apparently opposing requirements for transaxle fluid.

Plumbed Car Tests — With the evidence that transmission failure could occur quickly with EP-additive-containing lubricants even in ordinary driving, it was desirable to obtain more experience using various lubricants under these conditions. Since a transaxle was not available for fluid evaluation, it was necessary to continue using available current equipment to simulate transaxle operation.

The Research Laboratories garage cars were available for fluid testing. Also available were some Research Laboratories fuel-test cars used at the Proving Ground at Milford, Mich. The cars in both groups are driven in "transportation service" by company employees. Trips include home-to-work service as well as intercity trips. The amount of full-throttle acceleration to which these cars are subjected is probably greater than that encountered in average customer driving. Also, when in use for fuel evaluation, the fuel-test cars are no doubt driven harder than the average automobile is. On the other hand, operation of cars in both groups is generally confined to the northern part of the country, principally Michigan, and the operating temperatures are relatively low.

It would have been possible simply to install test fluids in the axles and transmissions of these cars and observe their performance as miles were ac-

[51] "Application of Wear Resistant Phosphate Coatings to Ferrous Surfaces," by M. B. Roosa. *Lubrication Engineering*, Vol. 6, June, 1950, pp. 117–121.

[52] Procedures for the tests included in this report may be obtained from the Fuels and Lubricants Department, Research Laboratories, General Motors Corp., 12 Mile and Mound Roads, Warren, Mich.

Table 2 — Effect of Phosphate Coating on Axle Hypoid Gears

Lubricant	Phosphate Coated	GMR No. 2 Gear Test			
		Gear Evaluation[a]		Noise	Remarks
		Ring	Pinion		
Proposed Transaxle Fluid (C)	Yes	2	3	Trace	
Proposed Transaxle Fluid (C)	No	5	5	Very noisy	Did not finish test
Proposed Transaxle Fluid (E)	Yes	2	2	Trace	
Proposed Transaxle Fluid (E)	No	5	5	Very noisy	Did not finish test

[a] Scale of 0 (no tooth surface distress) to 5 (very bad distress); 3 borderline.

Table 3 — Plumbed Car Test Results

Car	Transmission	Proposed Transaxle Fluid	Total Miles	Remarks
Buick	Dynaflow	O	14,435	Test discontinued, no failure
Buick	Dynaflow	P	15,868	Mechanical failure of plumbing circuit
Oldsmobile	Hydra-Matic (controlled-coupling)	Q	23,400	Mechanical failure of plumbing circuit
Pontiac	Hydra-Matic (controlled-coupling)	A	22,000	25,000-Mile General Durability Schedule — pinion bearing failure
Oldsmobile	Hydra-Matic (controlled-coupling)	A	19,000	Car is still in service

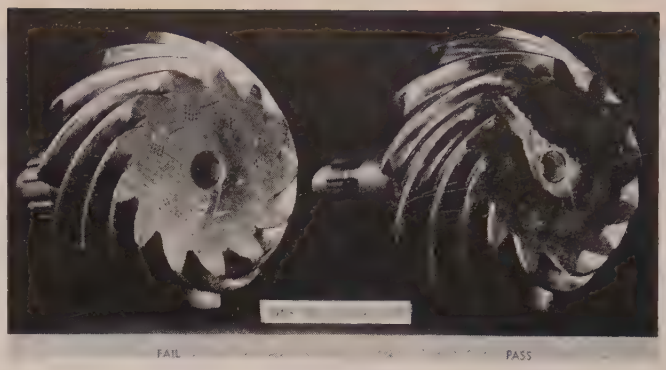


Fig. 1 — Hypoid gear test results showing effect of adding extreme pressure additive to automatic transmission fluid. Gear on left shows result of test run on automatic transmission fluid. Gear on right was tested with same automatic transmission fluid to which was added extreme pressure additive

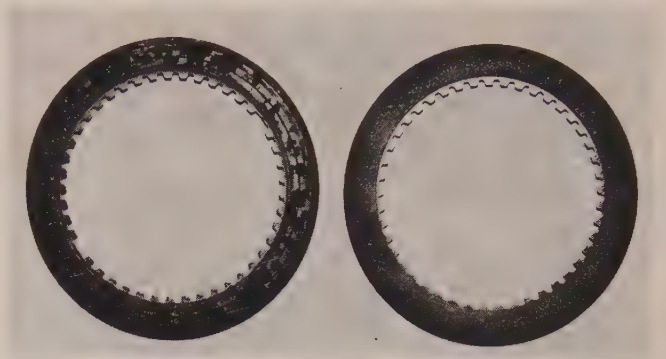


Fig. 2 — Effect of extreme pressure additive on clutch plate performance. Clutch plate on left shows adverse effects of extreme pressure agent in first proposed transaxle fluid tested. Clutch plate was operated for only 50 miles of ordinary driving in car. Clutch plate on right is still in good condition after 10,577 miles of operation in car on automatic transmission fluid

cumulated. However, it was recognized that in testing fluids in only the axle or transmission, certain interaction effects might have gone undetected. For example, there was some concern about the effect of the particles of phosphate coating wearing off the hypoid gears and causing malfunction of the transmission. Conversely, the oxidation severity level may be more severe in the transmission than in the axle, so that there might be an adverse effect on the EP quality of the lubricant. A simple means of checking for such possible interactions is to drain the axle and transmission at prescribed intervals, mix the fluid thoroughly, and re-install it in the two units.

Another method that was tried in a number of cars was that of plumbing the car to circulate the fluid between the axle and transmission. The advantage of this plumbed-car system is that the fluid in the transmission and axle is automatically and continually mixed. Table 3 shows some of the results obtained in cars equipped with this system. The test using Fluid O was discontinued because other work on this fluid showed that it was incompatible with clutch plates.

Tests run using Fluids P and Q resulted in axle gear failure due to an insufficient amount of lubricant in the unit, which was caused by mechanical failure of the plumbing system. The test using Fluid A, which was installed in a Pontiac car and run on the 25,000-Mile Proving Ground General Durability Schedule, was discontinued at 22,000 miles due to the failure of the differential front pinion bearing. At the end of this test, the transmission was in good condition and was exceptionally clean. The pinion bearing failure may have been caused by malfunction of the plumbing system. Another test is being run on Fluid A in an Oldsmobile with which 19,000 miles have been accumulated without difficulty.

As indicated, considerable difficulty was encountered in the plumbed car testing, such as pump, valve, and electrical failure. Therefore, it seems that it would be more practicable to conduct this sort of test using the draining-and-refilling procedure mentioned previously.

It is true that the plumbed car testing was plagued by difficulties associated with the plumbing; nevertheless, a considerable amount of valuable information has been obtained from this work. For example, the presence of phosphate-coating particles apparently does not cause any difficulty in the transmission. Also, it is possible to accumulate more than 20,000 miles without difficulty with a common fluid. Finally, the plumbed cars provided valuable shear stability data.

Shear Stability Studies — Fluid shear stability has long been recognized as a critical characteristic of automatic transmission fluid. In fact, at the present time serious consideration is being given to making the shear stability requirement more stringent

in the GM Automatic Transmission Fluid Specification. With a shear-sensitive fluid, an excessive permanent viscosity loss can occur with use. The resulting transmission fluid pressure loss can cause transmission malfunction and even mechanical damage. In addition, it is possible that greater leakage will occur with low-viscosity, sheared fluid. The resultant loss of fluid can also cause mechanical damage.

The shear stability characteristic takes on even more significance in transaxle fluid development work, because data obtained thus far indicate that the rear axle gears shear the fluid to a greater extent than the transmission. The higher shearing rate is probably due to the combination of high pressures and high sliding velocities between the hypoid gears.

Fig. 3 shows viscosity data obtained in plumbed cars for Fluids A, P, and Q. These data are compared with similar data obtained on an Automatic Transmission Fluid, Type A, used only in transmissions of Research Laboratories garage cars.

It should be noted that the curve for Fluid P represents only two points; it was arbitrarily drawn parallel to the curves for the other fluids.

An average viscosity loss of 1.8 centistokes occurred in 15,000 miles with the plumbed cars compared to 1.1 centistokes for the garage cars in a similar type of service. Furthermore, the curves show that the viscosity has apparently reached equilibrium in the garage cars at about 10,000 or 15,000 miles, whereas the viscosity is continuing to drop in the plumbed cars at 18,000 miles. A loss of 1.8 centistokes is considered excessive for automatic transmission fluids. However, no adverse effects were observed in the plumbed car work, possibly because no very-high-temperature operation was involved. Also, the average viscosity of the fluids used in the plumbed cars was 7.19 centistokes when new, which is somewhat higher than the minimum Type A fluid viscosity of 6.96 centistokes new.

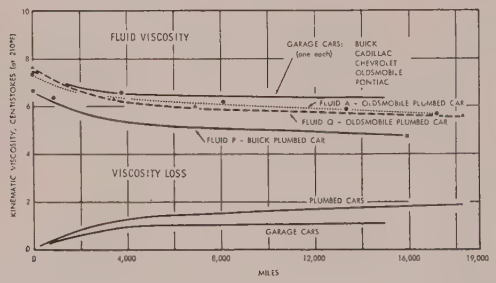


Fig. 3 — Effect of hypoid gears on change in viscosity (garage cars: fluid used in transmission only; plumbed cars: fluid circulated between transmission and axle, Table 3)

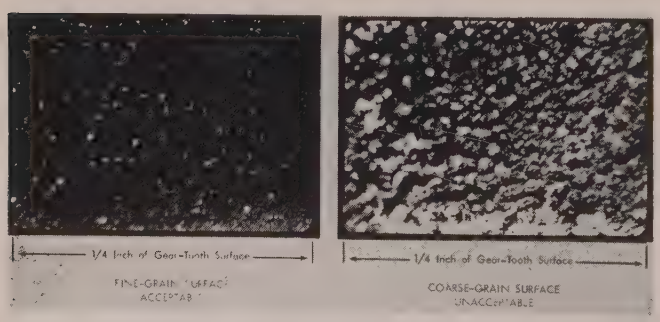


Fig. 4 — Variation of phosphate-coated gear surfaces

Tests Used for Evaluating Transaxle Fluid — The plumbed car tests showed that Dynaflow and Hydra-Matic transmissions will function satisfactorily for more than 15,000 miles when operated on a fluid used in common with the rear axle. These results were encouraging; the development of transaxle fluid no longer appeared so formidable as it did after the very early tests.

The plumbed car results are not very comforting to a transaxle engineer, however. From previous experience with automatic transmissions and rear axles he has learned that tests more severe than the plumbed car test must be employed to assure satisfactory operation of equipment in the customers' hands. He must apply tests to account for the "normal severe abuse" to which a certain percentage of the owners subject their cars. Some of the fluids which operated satisfactorily in the plumbed cars may not pass such tests. Accordingly, tests were chosen which have been used for many years by car division engineers to test transmissions, axles, and their lubricants. In addition, others which have proved to be useful in testing automatic transmission fluid were selected. In Table 4 these tests are listed together with the related requirements.

As discussed earlier, phosphate coating of hypoid gears is generally beneficial and only coated gears have been used in the tests reported here. The quality of the phosphate coating was found to be very important, however.

Fig. 4 shows the variation in the phosphate-coated gear tooth surface that was found among the gears used in the score testing. The fine-grain surface has been found to give good results. The coarse-grain surface gives the opposite effect. For example, fluids which passed the GMR No. 2 Gear Test with gears having the fine-grain surface failed very early in the test with gears having the coarse-grain surface.

These tests have shown that, if phosphate-coated

Table 4 — Tests Used for Evaluating Transaxle Fluid

Test	Requirement
Plumbed Car Test	Operation under customer-type driving conditions
GMR Hypoid Gear Test for Transaxle Fluid, No. 2	Hypoid gear tooth surface protection
	Hypoid gear frictional noise (reverse noise) prevention
Powerglide Transmission Inertia Cycling Test	Clutch-plate lubrication
Chevrolet Fade Test	Clutch-plate lubrication
Hydra-Matic Transmission Inertia Cycling Test	Clutch-plate lubrication
	Viscosity stability
	Oxidation resistance and thermal stability
	Seal compatibility
Proving Ground 5000-Mile LL Schedule	Lubrication of clutches and other transmission parts
Powerglide Oxidation Test (CRC Designation L-39) with and without air	Oxidation resistance and thermal stability
Tentative Procedure for Brookfield ATF Analysis	Low-temperature fluidity
Seal Dip Cycle Test	Seal compatibility
Total Immersion Test (ASTM D471-57T)	Seal compatibility

gears are used, it is essential to have a fine-grain coating on the gear tooth surface for evaluation of transaxle fluid EP quality. The tooth surfaces of new test gears should be carefully inspected in order to avoid very misleading test results.

Incidentally, it should be noted that a limited amount of test work has indicated that the CRC Reference Gear Oil performance level system may not apply to phosphate-coated gears.

Mineral-Oil-Base Fluids — Thirty-two mineral-oil-base fluids received from oil and additive suppliers have been evaluated. Of these fluids, only 11 passed the gear test and only 11 passed the transmission clutch plate tests. Furthermore, some fluids showed deficiencies in such properties as viscosity and oxidation stability. Of the 32 fluids, only the three listed in Table 5 passed both the gear and the clutch-plate tests. Even these three have certain undersirable characteristics and work is continuing in the effort to improve them.

Generally, the cost of these mineral-oil-base fluids is not expected to be substantially higher than that of current automatic transmission fluid. While it is recognized that fluid cost is a very important consideration, it is also recognized that other products, such as synthetic fluids, should be investigated even though they might be considerably more expensive.

Synthetic-Base Fluids — Some synthetic-base fluids have inherent properties which are obtained in mineral-oil-base fluids only through the use of additives. For example:

1. The viscosity index of some of these fluids is as high as 140. To obtain this 140 V.I. in a mineral-oil-base fluid, as much as 10% viscosity index improver must be added. Such addition gives rise to the problem of permanent viscosity loss due to shearing of this V.I. improver.

2. The oxidation resistance and thermal stability of some synthetic products with no additive treatment are as good as those of mineral-oil-base fluids containing additives.

3. Some synthetic products have antiwear properties superior to those of mineral oils of the same viscosity. This means that less additive may be required to obtain a given EP quality level.

It is because of these inherent properties that there is an interest in synthetic products despite the cost of these materials which ranges from $2 to as high as $30 per gallon, compared to an expected cost of less than $1 per gallon for a mineral-oil-base transaxle fluid.

Table 6 shows results of tests conducted on the 10 synthetic fluids which have been evaluated. As shown, only three of these fluids had sufficient gear lubricating qualities to pass the GMR No. 2 Gear Test. Further testing of these three fluids indicated deficiencies in other properties; specifically, viscosity at 210 F, oxidation resistance, and clutch-plate compatibility. None of the synthetic-base fluids tested compares favorably with the mineral-oil-base fluids given in Table 5 except, possibly, Fluid I, which gave a borderline result in the Chevrolet Fade Test.

Additive Evaluation Studies — Another approach to transaxle fluid development was the study of fluids compounded at the Research Laboratories. In this study, the effect of EP additive concentration was examined in some detail.

The first approach was the addition of small increments of EP additives to fully blended automatic transmission fluids and evaluation of these fluids for gear lubrication quality. Results of such tests using one additive, di(n-octyl) phosphite (DOP), are shown in Table 7. With concentrations of 2% or less, no EP benefit could be observed. Maximum benefit was obtained at concentrations of 4, 6, and 8%.

Unfortunately, fully blended automatic transmission fluids contain additives, such as antisquawk and oxidation inhibitor additives, which also impart mild EP properties and thus tend to mask the incremental effects of the EP additive being evaluated. For instance as shown in Table 7, borderline results were obtained in the test on the proposed automatic transmission fluid containing no supplementary EP additive. Accordingly, an additive carrier was blended that would permit the observation of small differences in the effectiveness of various EP additive concentrations. This carrier was made up to give viscosities in the automatic transmission fluid range. No additives were included that are believed to impart antiwear or EP quality to a fluid.

The additive carrier formulation and viscosity are as follows:

Formulation	Viscosity
95.5% (by weight) ATF base oil	7.19 cs at 210 F
4.5% (by weight) V.I. improver	36.37 cs at 100 F
20 ppm foam inhibitor	45,550 cp at –40 F

To establish that the carrier contained a minimum of antiwear or antiscore properties, two GMR

Table 5 — Mineral-Oil-Base Fluids

Fluid	GMR No. 2 Gear Test	Chevrolet Fade Test	Proving Ground 5000-Mile 7/11 Schedule	Powerglide Inertia Cycling	25,000-Mile General Durability[a]	Powerglide Oxidation[d]	Viscosity-Temperature Characteristics
A	Pass	Pass	3 tests: 1 pass, 1 borderline, 1 fail at 4800 miles	3 tests: Pass	Pinion brg.[b] failure at 21,000 miles	Condition 1: pass Condition 2: fail	Pass
B	Pass	Pass	—	—	—	—	—
C	Pass	Pass	—	2 tests: Pass	Pass[c]	Condition 1: pass Condition 2: fail	Pass

[a] The 25,000-Mile General Durability Schedule is run at GM Proving Ground as a final test of car performance. It includes a wide range of driving conditions.

[b] Plumbed Car Test using a Pontiac—pinion bearing failure may have been due to oil starvation resulting from circulation system failure.

[c] This test was run 35,000 miles using a Chevrolet.

[d] Test run under both conditions 1 and 2.

Table 6 — Synthetic-Base Fluid Test Results

Lubricant	Type	GMR No. 2 Gear Test: Gear Evaluation Ring	GMR No. 2 Gear Test: Gear Evaluation Pinion	GMR No. 2 Gear Test: Noise	Chevrolet Fade Test	Remarks
E (turboprop oil)[a]	Diester	3	4	Bad backup		
F (helicopter gear lubricant)	Polyglycol	0	1	Light whine	Pass	Failed viscosity, too low at 210 F
G (proposed transaxle fluid)	Polyglycol	2	2	Trace	Pass	Failed Powerglide Oxidation Test; borderline failure in 2 Powerglide Transmission Inertia Cycling Tests
H (base oil)	Diester	4	4	Medium		
I (proposed transaxle fluid)	Diester	0	1	Trace	Borderline	
J (turboprop oil)	Diester	5	5	Low whine		
K (base oil)	Diester	4	4			
L (base oil)	Diester	5	5	Trace growl		
M (proposed transaxle fluid)	Diester	4	4	Light growl		
N (base oil)	Polyglycol	3	3	Very heavy noise		Test stopped due to noise

[a] Approximately one-half of the GMR No. 2 Gear Test was run.

Table 7 — Effect of Additive Concentration on Hypoid Gear Performance

Base Fluid	Extreme Pressure Additive	GMR No. 2 Gear Test: Conc. Wt. %	Gear Evaluation Ring	Gear Evaluation Pinion	Noise
Proposed ATF[a]	None	0	1	3	Heavy reverse
Proposed ATF	Di(n-octyl)phosphite	1	2	3	Medium whine
Proposed ATF	Di(n-octyl)phosphite	1¼	2	3	Trace
Proposed ATF	Di(n-octyl)phosphite	1½	1	3	Trace whine
Proposed ATF	Di(n-octyl)phosphite	2	2	2	Trace
Proposed ATF	Di(n-octyl)phosphite	4	0	1	Trace
Proposed ATF	Di(n-octyl)phosphite	6	0	0	Trace whine
Proposed ATF	Di(n-octyl)phosphite	8	0	0	Light whine
Additive Carrier[b]	None	0	5	5	Medium
Additive Carrier	Treated tri(2-chloro-1-methylethyl)phosphite	¼	1	2	None
Additive Carrier	Treated tri(2-chloro-1-methylethyl)phosphite	½	0	0	None
Additive Carrier	Treated tri(2-chloro-1-methylethyl)phosphite	1	0	0	None
Additive Carrier	Treated tri(2-chloro-1-methylethyl)phosphite	2	0	0	Trace
Additive Carrier	Di(n-octyl)phosphite	2	0	2	Trace

[a] Automatic transmission fluid.
[b] Additive carrier: 95% (by weight) ATF base stock
4.5 % (by weight) V.I. improver
20 ppm foam inhibitor.

No. 2 Gear Tests were run using the carrier as the test lubricant. In both tests the gears were badly scored.

Results obtained using the additive carrier and treated tri phosphite are shown in Table 7. As little as ¼% of this additive gave satisfactory results and concentrations of as much as 2% gave similar results.

It should be emphasized that the research additive carrier plus extreme pressure additive does not represent a finished transaxle fluid. Other additives are required to impart additional properties such as oxidation resistance, antisquawk quality, detergency, and the like. It is realized that this procedure for evaluating additives alone may give rise to an additive interaction problem. That is, additives evaluated individually do not necessarily react in the same manner when evaluated with other additives. Consequently, once an additive has been proved effective in the carrier, it must be re-evaluated in a fully blended transaxle fluid.

There is evidence of additive interaction in the test with 2% DOP in the proposed automatic transmission fluid. Table 7 shows that the test run with 2% DOP in the additive carrier gave a "0" ring gear rating compared to the "2" rating for the 2% DOP – proposed automatic transmission fluid test. In the latter test the automatic transmission fluid additives apparently interfered with the action of the DOP.

Additional additive evaluation studies are summarized in Table 8. Extreme pressure additives were evaluated in fully blended automatic transmission fluid and the additive carrier previously discussed. In addition, test results are shown for three experimental transaxle fluids blended at the Research Laboratories.

As shown in the table, only seven of the 20 gear tests conducted proved satisfactory. Four of the additives, di(n-butyl)phosphite, di(n-octyl)phosphite, treated tri(2-chloro-1-methylethyl)phosphite, and GMR No. 1313 proved to be incompatible with the Powerglide clutch plates. However, because of the favorable gear results obtained with lower concentration of treated tri phosphite, as shown in Table 7, further fluid-clutch plate compatibility work will be conducted using this additive. One additive, di(n-octadecyl)phosphite, proved unsatisfactory due to insolubility at temperatures below 10 F. Therefore, only three of the additives evaluated showed sufficient promise to warrant further testing. These were:

1. Treated tri(2-chloro-1-methylethyl)phosphite
2. Di(n-dodecyl)phosphate
3. Di(n-dodecyl)phosphite

It is of interest to review the background of one of these additives, di(n-dodecyl)phosphite. Related extreme pressure agents were first evaluated

Table 8 — Additive Evaluation Studies

Base Fluid	Additive	Concentration by Weight %	Extreme-Pressure Quality GMR No. 2 Gear Test			Clutch Plate Compatibility			
			Gear Rating		Noise Rating	Powerglide Transmission Inertia Cycling Test	Hydra-Matic Transmission Inertia Cycling Test	Proving Ground 7/11 Schedule	Chevrolet Fade Test
			Ring	Pinion					
Proposed ATF[a]	—	0	1	3	Heavy reverse	—	—	—	—
Additive Carrier[b]	—	0	5	5	Medium	—	—	—	—
Proposed ATF	Di(n-octyl)phosphite	2	2	2	Trace	—	Pass	Fail	Pass
Proposed ATF	Di(2-ethylbutyl)phosphite	2	5	5	Heavy noise	—	—	—	—
Proposed ATF	Di(2-butyloctyl)phosphite	1½	2	4	Medium	—	—	—	—
Proposed ATF	Di(2-ethylhexyl)phosphite	2	5	5	Heavy noise	—	—	—	—
Additive Carrier	Di(n-octyl)phosphite	2	0	2	Trace	—	—	—	—
Additive Carrier	Tri(n-dodecyl)trithiophosphite	2	5	5	Very noisy	—	—	—	—
Additive Carrier	Di(n-dodecyl)phosphite	2	0	2	Trace	—	—	—	—
Additive Carrier	Di(n-octadecyl)phosphite	2	0	1	None	—	—	—	—
Additive Carrier	Treated tri(2-chloro-1-methylethyl)phosphite	2	0	0	Trace	—	—	—	—
Additive Carrier	Di(n-dodecyl)dithiophosphite	2	4	5	Heavy reverse	—	—	—	—
ATF I	Dibenzyldisulfide	2	3	4	Heavy growl	—	—	—	—
ATF I	Lead di(n-amyl)dithiocarbamate	3	5	5	Growl	—	—	—	—
ATF I	Zinc dialkyldithiophosphate	3	3	4	Heavy reverse	—	—	—	—
ATF I	Chlorinated wax	2	5	5	Heavy whine	—	—	—	—
ATF II	Di(2-chloroethyl)vinylphosphonate	2	4	4	Medium	—	—	—	—
ATF II	Di(n-butyl)phosphite	2	—	—	—	3 tests: fail	—	Fail: 1099 miles	—
ATF II	Di(n-dodecyl)phosphite	2	—	—	—	2 tests: borderline	—	—	—
ATF III	Treated tri(2-chloro-1-methylethyl)phosphite	1	—	—	—	—	—	—	Fail
GMR Transaxle Blend No. 1313	—	—	0	2	Trace	Fail	—	—	—
GMR Transaxle Blend No. 1319	—	—	0	2	Trace	—	—	—	—
GMR Transaxle Blend No. 1320	—	—	3	4	Trace	—	—	—	—

[a] Automatic transmission fluid.
[b] Additive carrier: 95.5% (by weight) ATF base stock
4.5% (by weight) V.I. improver
20 parts per million foam inhibitor

for gear lubrication effectiveness and fluid-clutch plate compatibility. These agents and their chemical formulas are:

Di(n-butyl)phosphite $(C_4H_9O)_2P(O)H$
Di(n-octyl)phosphite $(C_8H_{17}O)_2P(O)H$
Di(n-dodecyl)phosphite $(C_{12}H_{25}O)_2P(O)H$

All three were found to have satisfactory gear lubricating qualities. However, they differed significantly from one another with respect to fluid-clutch plate compatibility. For example, in concentrations of 2%, di(n-butyl)phosphite flaked clutch plates severely, di(n-octyl)phosphite somewhat less severely and di(n-dodecyl)phosphite flaked clutch plates the least. In other words, clutch-plate flaking decreased as carbon chain length increased. Longer carbon chain compounds of this class such as di(n-octadecyl)phosphite, $(C_{18}H_{35}O)_2P(O)H$, appear to be impractical because of solubility difficulties.

Results of tests conducted on three transaxle fluids blended in the Research Laboratories are also shown in Table 8. Only one of these, GMR No. 1319, appears promising.

Summary

Hypoid gear lubricant has been under development for about 34 years and automatic transmission fluid for about 23 years. General Motors has been developing transaxle fluid for the past several years. In this development they have had the cooperation of oil and additive suppliers.

Plumbed car tests of up to 23,000 miles indicated that phosphate-coating debris from the hypoid gears will cause no transmission difficulty. However, this work showed that viscosity reduction may be expected to be greater in the transaxle than in the automatic transmission. Fluids which are satisfactory in plumbed car tests may not be satisfactory in other full-scale transmission and axle tests.

More stringent tests used in transaxle fluid evaluation included a hypoid gear bump test and automatic transmission clutch-plate inertia cycling tests. In addition, bench tests have been used to measure viscosity and seal compatibility. Many fluids have been evaluated including 32 mineral-oil-base fluids and 10 synthetic-base fluids.

In addition, numerous additive-base stock combinations have been tried. The results of the initial work reported here indicated that three mineral-oil-base fluids and one synthetic-base fluid obtained from suppliers, and one additive-base stock combination blended at the Research Laboratories merited further investigation. It must be remembered, of

course, that final proof-testing of any transaxle fluid must be performed in actual transaxle equipment.

The conclusion drawn from this early work was that apparently it is feasible to operate a transaxle on a single fluid. Satisfying the requirements of both the hypoid gears and the transmission clutch plates appears to be one of the more difficult problems in developing a fluid. While the development of satisfactory transaxle fluid has met with some success, considerable work remains to be done in the matching of transaxles and fluids to achieve optimum results.

DISCUSSION

Gear Design Also A Factor In Transaxle Fluid Performance

— D. F. Greene

Esso Research and Engineering Co.

OUR WORK at the Esso Co. has corroborated many of the points raised by the authors. We agree that the essence of the transaxle lubricant problem is to get the level of EP activity needed for gear protection while simultaneously getting the good thermal and oxidative stability required for transmission performance.

Fig. A illustrates the sort of thing one runs into in trying to hit these two targets simultaneously. Here are photographs of transmission and rear axle parts taken from a conventional 1959 automobile that had been operated for about 25,000 miles of typical motorist driving with the same experimental lubricant in the axle and in the automatic transmission. (The axle had phosphated gears.) These units are in excellent condition, with no mechanical distress and freedom from deposits and corrosion. The thrust washers and clutch plates were in excellent condition and the shifting and other operational characteristics of the car appeared entirely normal.

Based on this good car performance, one is encouraged about the prospects for solving this fluid formulation problem. The situation brightens further when we realize that this experimental fluid also successfully passes the severe rear axle shock test, and also meets the Bench Test requirements for Type A automatic transmission fluids. On the other hand, however, this fluid allows corrosion of the babbitt metal and copper alloy parts in the CRC L-39 transmission oxidation test — one of the requirements for transaxle fluids that have been mentioned by the authors — although no corrosion was encountered in the 25,000-mile test in the 1959 automobile. We think that this is a consequence of the higher oil temperature in the L-39 test — 275 F versus temperatures of about 200 F in the car.

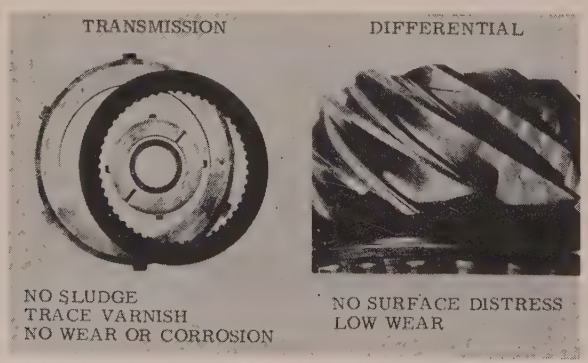


Fig. A — Experimental transaxle fluid A, 24,000 miles (70% stop and go, 30% high speed)

We also find that this same experimental fluid permits considerable wear of the axle pinion gear when tested in the transaxle-equipped Corvair car. No such wear was encountered in the conventional 1959 car, as we observed before. Now, this may be a consequence of the different gear loadings or possibly other gear design differences between these two car types.

Two questions suggest themselves at this point, therefore:

1. Will performance requirements at 275 F (the L-39 test temperature) continue, or might it be possible to alter transmission cooling characteristics and reduce transaxle bulk oil temperatures closer to the 200 F performance level?

2. The fact that a lubricant will perform excellently for extended service in one type of axle, but not in another, suggests that there may be room for further investigations in the gear design area.

I'd like to comment next on two other aspects of the transaxle lubricants picture. The first point concerns the effect of the quality of the phosphated coating on the gears (Fig. B). We agree with the authors' observation that a fine-grain uniform coating leads to better results than a coarse-grain or spotty coating. These photographs are Faxfilm impressions of the gear surfaces taken before the rear axle bump test was run. The coarse-grain phosphated surface on the right gave a failure in the rear axle shock test, while the fine grain coating on the left gave a good pass on the same lubricant in the shock test.

This illustrates the importance of selecting phosphated gears for lubricant research work, and may also indicate an area for further research by the gear makers. It also raises the question of whether the instance of poor performance that the authors encountered on a commercial SAE 90 gear oil (using phosphated gears) might be associated with the phosphate coating itself, as well as possibly being associated with an interaction between the EP agents and the phosphate coating, as they have suggested.

My last point concerns the relative shearing actions of the transmission and the axle. Our experience parallels that of the authors in indicating that the axle does more shearing than the transmission (Fig. C). The height of these bar graphs represent the viscosity of the transaxle fluid, taking the new oil 100%. After 25,000 miles in the conventional 1959 car of Fig. A, the oil viscosity is reduced

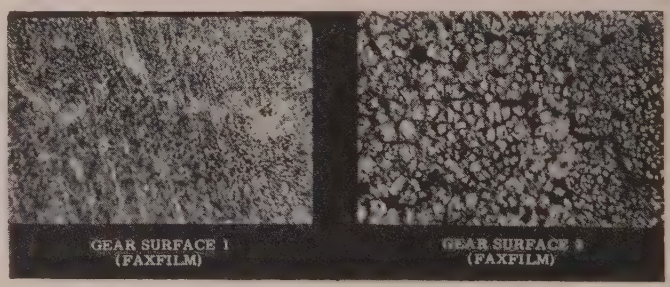


Fig. B — Experimental transaxle fluid B, axle "shock" tests

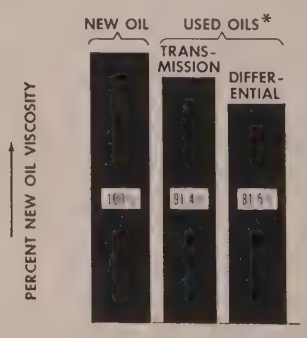


Fig. C — Experimental transaxle fluid A, per cent viscosity (*24,000 miles, 70% stop and go, 30% high speed)

by about 9% in the transmission. In the differential, we've lost more than that, or about 19%. We have observed a similar pattern in the Corvair car. Again, the axle gave more shearing than the automatic transmission. Possibly the situation can be improved upon by the use of polymers with super-shear stability. Possibly also, further experience may show that some gear designs are less severe than others in this respect. For example, spiral bevel gears might give less shearing than the hypoid gear type.

Traxle Fluids Performed Well In 1959 Oldsmobile Road Tests

— N. V. Messina and L. W. Manley
Socony Mobil Oil Co., Inc.

SOME OF us, who have been in the lubricant formulation business for many years, were quite skeptical at the outset of this development that the goals established for a combined automatic transmission fluid and hypoid gear lubricant colud be satisfactorily achieved. Based on the information presented in this paper plus our own laboratory experience, we are pleasantly surprised and greatly gratified at the progress that has been made. This progress has indeed been the result of close cooperation between the automotive builders and the lubricant or additive suppliers. There are still serious problems to overcome before the development is complete; however, the outlook is encouraging based on the progress to date.

We were pleased to note the satisfactory results reported in the preliminary road testing of experimental traxle fluids. We also have some experience in this regard which may be of interest. Using a fluid, which looked promising at the time but which in the light of current requirements could not be considered satisfactory with respect to laboratory performance testing, a road test was conducted in the six passenger cars shown in Table A. The 1959 Oldsmobile was driven over a prescribed turnpike-city driving course for 25,000 miles by professional drivers. The other five cars were employee-owned and operated in normal mixed suburban-turnpike service. The transmissions were overhauled at the start of the test, and new rear axle assemblies with phosphate-coated gears were installed in each car. The same oil was used in both automatic transmission and rear axle. It was drained from both transmission and rear at 1000-mile intervals, thoroughly intermixed and then put back in each unit. This afforded a condition somewhat comparable to what might be expected in combined transmission-rear axle service.

The inspections made at the conclusion of the tests showed all transmissions were operating normally and smoothly, all parts were in good condition with no evidence of clutch plate distress, and with only trace amounts of deposits (probably phosphate-coating debris). The excellent condition of the transmission and rear axle parts from the 25,000 Oldsmobile test is shown in Figs. D–I. The used oil showed a substantial drop in viscosity from the initial 51 SUS which agrees with the authors' observations that V.I. improved oils are very severely sheared in rear axle service.

Obviously, these road tests were not conducted under the severest conditions to which automotive equipment can be subjected. However, the results do indicate satisfactory performance in a normal type of operation even with an early model, now obsolete experimental traxle oil.

Our development program on traxle fluids confirms sev-

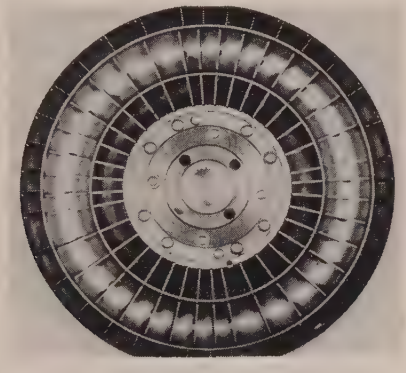


Fig. D

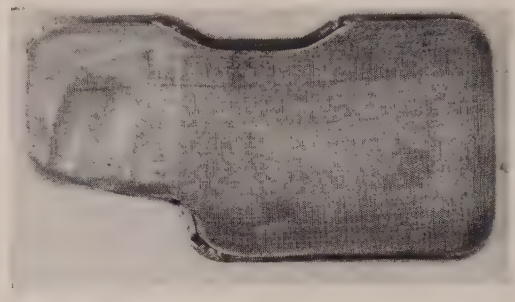


Fig. E

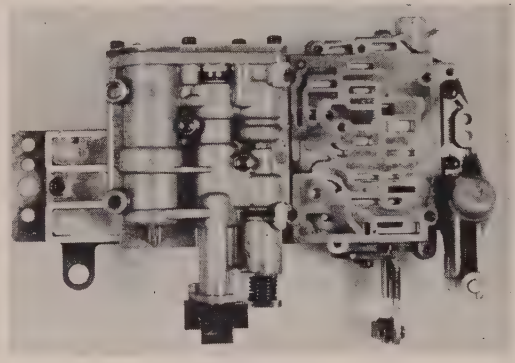


Fig. F

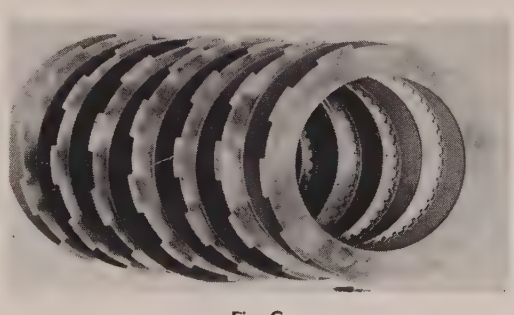


Fig. G

Table A — Experimental Traxle Oil Road Test

	1959 Oldsmobile	1958 Oldsmobile	1958 Oldsmobile	1958 Chevrolet	1958 Chevrolet	1958 Mercury
Test Miles	25,000	12,400	11,600	15,000	15,000	15,000
Final Used Oil SUS at 210 F	44.4	44.4	—	44.6	43.4	44.3
% Pent. Insolubles	0.03	0.03	—	0.03	0.01	0.03
Final Transmission Operation	Normal	Normal	Normal	Normal	Normal	Normal
Cleanliness	Lt. Black Wipeable Film Overall					
Final Rear Axle Cleanliness	Clean	Clean	Clean	Clean	Clean	Clean
Gear Condition	Satisfactory					

NOTE: Every 1000 miles rear and transmission fluids were drained, mixed well, and put back in transmission and rear axle.

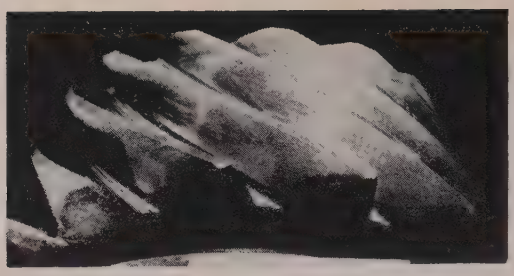


Fig. H

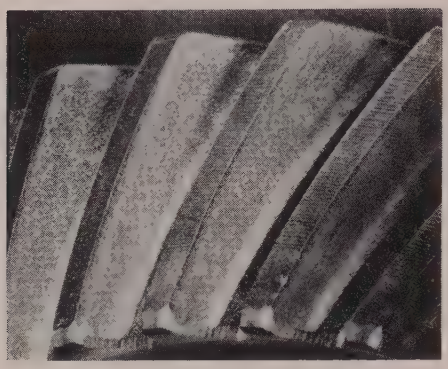


Fig. I

eral of the observations reported in the paper. We particularly agree that the technology of lubrication of phosphate-coated gears is quite different than untreated gears and also that the nature of the phosphate coating has a profound effect on performance. In addition, we have observed that some additives normally present in automatic transmission fluids greatly depreciate the effectiveness of E. P. additives.

There is the need for simpler, less costly, more repeatable, more standardized and perhaps, more realistic evaluation procedures. The use of the Brookfield viscometer as a replacement for a full scale vehicle cold room test cited by Mr. Hunstad is a superb example of what can be accomplished in this direction. The work of CRC in standardizing the Powerglide Oxidation Test is also a step forward. However, in the field of automatic transmissions and traxles there is evolving an ever-increasing number of complex evaluation tests, so as "to confound even the most elect." In order to improve this situation we indeed need a concerted, cooperative effort on the part of automotive builders and lubricant suppliers.

Fluids' Restrictive Requirements Should be Re-evaluated

— William H. Kite, Jr.
Texaco Inc.

THE AUTHORS imply that, with the transaxle incorporating the automatic transmission, the problem is to take an automatic transmission fluid and reformulate it to produce a multipurpose transaxle fluid. We certainly agree that this was the most obvious approach and definitely the one which deserved first trial. However, in view of subsequent experience and knowledge, we question whether the problem has not evolved into one which demands more basic consideration and research by both the automotive and petroleum industries.

The requirements which have been placed upon automatic transmission fluids throughout the years have increased progressively in number and severity and, in some instances, would seem to have exceeded the actual demands of the automatic transmissions themselves. In most cases, these automatic transmission fluid requirements now have reached almost formidable proportions. On the other hand, the development of gear oils for modern differentials is no easy task in itself. When the requirements of these gear oils are added to those already existing for automatic transmission fluids, the combined problem becomes almost astronomical.

It would seem that the whole subject deserves a close examination by all of us with the objective of attempting to eliminate as many of the restrictive requirements of the separate and combination fluids as practically feasible. Possibly, a close scrutiny of the differential oil and automatic transmission fluid specifications, particularly the latter, in the light of necessity and actual field experience is in order. Such an examination might well uncover restrictive requirements which, though desirable, are not necessary in the case of the individual fluids and which, when relaxed or eliminated, would not adversely affect transaxle performance. Even a seemingly minor such relaxation or elimination sometimes can expand vastly the possibilities for fluid development.

A simultaneous approach would consist of a close examination of the mechanical design of the transaxle components. The purpose of such an examination would be to uncover possible methods of design change and resultant relaxation of fluid requirements. The authors point out that steps already have been made in this direction through the phosphating of the differential gears and the increase in clutch plate area in the automatic transmission. We definitely concur in this type of approach and are certain that continued efforts along this line will simplify and benefit fluid development.

Authors' Closure To Discussions

MR. GREENE asked whether the unsatisfactory results obtained with the commercial SAE 90 grade gear lubricant using phosphate-coated gears could have been due to the poor quality of the coating. While it was believed that the coating was satisfactory, a close examination of the gears was made following Mr. Greene's suggesting that it might not be. It was found that the coating was *not* of the type that we have learned is required for satisfactory performance. Therefore, further tests will be required to determine whether the SAE 90 grade gear lubricant will be satisfactory with properly coated gears.

The other comments made by the discussers are generally related to one of the following: (1) test requirements, and (2) transaxle materials. Pleas have been made both for a reduction in test requirement severity and for a selection of transaxle materials that will reduce the lubricant requirement level.

We agree that such changes would make the transaxle lubricant development job easier. But there are two important considerations which apply to changes in either of the foregoing. First, it must be remembered that the ultimate objective is the development of a transaxle and fluid combination that will provide satisfactory operation at a minimum cost. Any change in transaxle cost must be weighed against the change in price of the lubricant. For example, an oil cooler could be provided to maintain a moderate maximum fluid temperature. However, the cost of such a cooler might more than offset any saving in fluid cost.

The other fact that must be borne in mind is that there is no common-sump transaxle commercially available. Until experience is gained with such units in the hands of the public, the true requirements of the transaxle will remain unknown. Accordingly, the current efforts to develop a transaxle fluid will, at best, result only in an approximation of the required material. It seems prudent, therefore, to retain the requirements based on past experience with conventional units until more experience is gained with common-sump transaxles.

Radioactive Cylinders

— a tool for wear research

W. C. Arnold and V. T. Stonehocker
Fairbanks, Morse and Co.

W. J. Braun and D. N. Sunderman
Battelle Memorial Institute

THIS PAPER describes an investigation into the use of radioactive cylinders in wear research. Battelle Institute and Fairbanks, Morse agree that the method has potential as a research tool.

The effect on cylinder wear experienced with variables such as starting, idle or cold temperature operation, sudden load changes, speed, torque, and such, can be determined easily and relatively inexpensively.

The authors describe the techniques and precautions taken during the testing.*

TODAY'S rapid technical growth has made it possible for all industry to utilize the knowledge gained in atomic engineering for a myriad of purposes. One such use is to ascertain quickly rates of wear of power parts in diesel engines. This, of course, is a subject of major importance to the entire diesel industry, especially in this era of economy fuel oils.

At the end of World War II, the use of radioactive isotopes in measuring engine piston-ring wear rates was contemplated by most of the diesel-engine and fuel-and-lubricant industries due to the evolution of the nuclear research reactor. Early reactor designs, such as is exemplified by the aircooled and graphite-moderated X-10 reactor at the Oak Ridge National Laboratory, have been widely used for producing these materials. A limitation of that family of reactors is the size of the specimens which can be irradiated. Small, uncomplicated engine parts, such as piston rings, valve guides, and bearing shells, can be accommodated. Larger items, including diesel-engine cylinder liners, could not be suitably radioactivated.

From the correlation of a great number of wear tests carried out over many years in the Fairbanks, Morse and Co. engine laboratory, it has been concluded that ring wear data bear little correlation to wear rates experienced in cylinder liners, at least for this company's engines. Obviously, cylinder wear is more important to the engine manufacturer and user than ring wear, because cylinders are more expensive to replace than piston rings. Conventional cylinder wear tests are very time consuming, since many operating hours must be accumulated to obtain accurately measurable wear increments. For this reason, a better and faster means of determining cylinder wear in engines was desired. As the first techniques for wear research with the use of radioisotopes became known, the then available equipment and techniques were carefully studied. It was hoped that a reliable and economical method to investigate the various phases of engine operation affecting cylinder wear might be developed. Such a program would have been impractical using

* Paper presented at SAE Annual Meeting, Detroit, Jan. 13, 1960.

Fig. 1 — Cylinder liner with external shield of cadmium and and internal shield of boral alloy

Fig. 3 — Placement of capsule in reactor

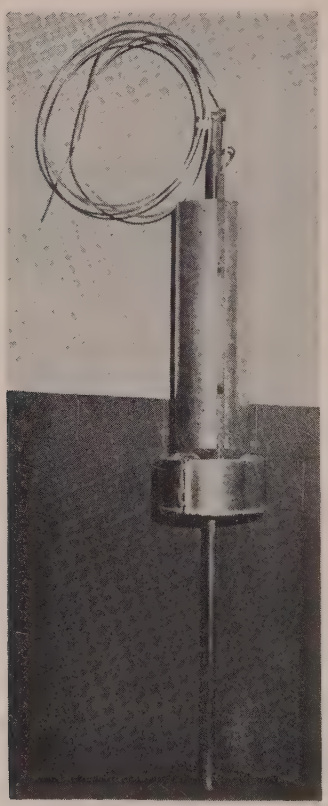


Fig. 2 — Aluminum capsule shielding liner during irradiation

conventional means of wear determination because of the time, money, and uncertainties involved in the long drawn-out programming required. Since at that time no feasible technique could be determined, the study was dropped.

During the ensuing years, the development of pool-type water-cooled and moderated reactors made possible the irradiation of larger sized specimens, allowing many companies to use the activated piston-ring technique in large bore engines. In 1957, Fairbanks, Morse and Co. felt that it would now be worthwhile to restudy the feasibility of irradiating a cylinder for use in wear studies. Of those approached, Battelle Memorial Institute was the organization which had both the physical facilities required plus the confidence of their technical staff that such a development project had a good possibility of success. A cooperative project was initiated early in 1958.

Since the area of maximum wear in our opposed piston engine cylinders is at the inner-dead-center end of the compression ring travel, our first request was to have a cylinder liner irradiated at this point to a high level of activity. If a technique could be developed for measuring wear at this point, it would be quite possible to shift later test work to other areas of the cylinder bore.

Another request made was to trace all of the irradiated cylinder bore material debris, since no known factor could be applied to determine what percentage would go to the lubricating oil sump for detection and what percentage would be dis-

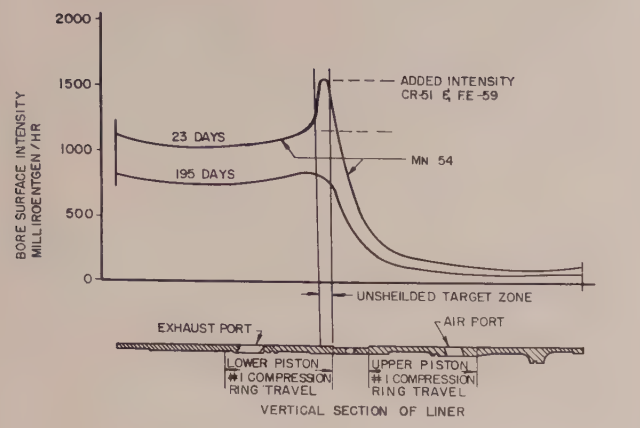


Fig. 4 — Radiation profile of liner

charged with the exhaust gases. A detection device for debris carried by the exhaust gases was a necessity. A third request was that the actual cylinder be irradiated rather than an inserted sleeve, due to the belief that the wear rate of the sleeve would not be representative of that experienced with the parent cylinder material.

Radioactive tracers have never been employed for measuring cylinder wear in unmodified cylinder surfaces of a large-capacity diesel. This fact necessitated the development of several important techniques coincident with this program.

Procedures were devised to determine the feasibility of radioactivating a specific area of a test liner. It was necessary to determine experimental parameters which woud result in the production of a usefully radioactivated cylinder.

Mechanical techniques were developed to permit the manipulation of a specimen as massive as the nuclear reactor core itself, which also would not result in a hazard to safe reactor operation.

Handling and installation techniques were developed which resulted in the minimum radiation hazard to personnel associated with the program.

Radioassay techniques were developed which permitted the precise analysis of fractions of wear product debris collected from lubricating oil and exhaust gas systems.

Irradiation Techniques

Prior to the irradiation of the cylinder liner, three short-term preliminary irradiations were conducted. These preliminary irradiations were designed to provide information which would serve as a guide for the radioactivation of a cylinder liner. They consisted of, first, the radioactivation and study of a typical piece of cylinder-liner metal; second, the irradiation and evaluation of candidate thermal-nergy neutron absorbers; and third, the irradiation of a section of cylinder liner fitted with a selected combination of neutron absorbers.

Cylinder Liner Irradiation

A 160-lb 42 in. long cylinder liner designed for use in a Fairbanks, Morse and Co. 8⅛-in. bore by 10-in. stroke opposed piston diesel engine, was selected for irradiation. The liner was provided with an external shield of cadmium and an internal shield of boral alloy in the manner illustrated in Fig. 1. A 1½ in. wide unshielded target band is visible at the position corresponding to the contact zone of the top two compression rings of the lower piston when at its inner-dead-center position. The wires visible about the circumference of that zone are aluminum-cobalt dosimeter wires used for recording the integrated thermal neutron flux incident to the target zone.

The shielded cylinder liner was installed in the watertight aluminum capsule illustrated in Fig. 2. The aluminum tube, later removed, was employed to pressurize the capsule with helium during irradiation. The gas served to prevent water leakage into the capsule during irradiation in the advent of a leak. The nuclear properties of helium are such that it does not become radioactive, thus minimizing the hazard to reactor personnel in case of capsule failure.

The helium-pressurized cylinder-liner capsule was irradiated for 4 weeks. Fig. 3 illustrates the placement of the capsule in the reactor and the drive mechanism which was employed to rotate the encapsulated specimen at a speed of 1 rpm. Capsule rotation resulted in an extremely even irradiation.

Fig. 4 illustrates the approximate radiation profile through a vertical section of the cylinder liner. This approximation was determined from the dosimeter wires and from scanning the liner bore with radiation detection equipment. The top curve illustrates, at 23 days after irradiation, the more radioactive target band produced by the presence of radioactive iron and chromium isotopes. The approximation at 195 days shows the band to have merged into the radiation level of the manganese because of the faster decay rate of the iron and chromium isotopes.

The following evaluation was made of the radioactive cylinder liner:

1. The attempt to radioactivate a liner in a small, specific zone to a comparatively high level of specific radioactivity was successful to a limit extent. With the employment of elaborate radiation detection equipment during analysis it would have been possible to discriminate wear occurring in that zone. However, it was not attempted in this program.
2. Manganese-54 was produced in useful quantities in the entire lower half of the liner. The 300-day half life of that radioisotope resulted in a liner with a useful radioactive life of years compared with that of months normally present in piston rings which are normally activated in a manner resulting in the production of shorter lived iron-59 and chromium-51.
3. The desirability for capsule rotation was confirmed by the evenly radioactived cylinder liner.

Handling and Installation Techniques

For this study, it was Battelle's responsibility to develop techniqes for handling and installing the radioactive cylinder liner in a Fairbanks, Morse diesel test engine. A careful evaluation was made of all operations which would be involved. This evaluation was then interpreted in terms of the radiation hazard presented by the radioactive liner. Han-

dling and installation procedures were considered in order to develop procedures which would minimize the radiation hazard to all persons who would be exposed.

The specially crated lead-shielded liner was transported from the reactor site at Columbus, Ohio to Beloit, Wis., by truck. Operations involving jacketing, water testing, honing, and bore checking were all performed by production personnel at the normal production station used for each operation. Fig. 5 illustrates the liner jacket installation. During these operations the radiation hazard created in the area of the liner was evaluated with radiation survey meters. All personnel concerned with liner handling in any way were provided with radiation-sensitive film badges for recording cumulative dosage, and direct-reading pocket dosimeters. All preinstallation handling operations were completed in 3 hr, and a maximum radiation dose of 80 milliroentgens (mr) was measured by the film badge for one operator. Other personnel received negligible dosages. Care was taken to insure the freedom of all areas and equipment from radioactive contamination.

Installation of the radioactive liner and engine buildup were completed in three 8-hr shifts. Two operators plus a supervisor were used for each shift. Film badges, dosimeters, and radiation-detection instruments were employed as safety measures. A Battelle representative was present at all times to supervise the handling of the radioactive liner and to make certain that radiation safety measures were properly executed. Radiation doses measured with pocket dosimeters ranged from 35 to 80 mr. No one involved in engine buildup received more than the permissible dose of 300 mr per week. Figs. 6 and 7 illustrate the engine buildup.

Following a normal buildup the engine and its immediate working area were evaluated as a radiation hazard. A suitable area was roped off about the engine at a sufficient distance to limit the radiation intensity beyond the designated zone to less than 2 mr per hour. Signs warning of the radiation hazard were posted to designate that zone. All personnel permitted inside the posted area were provided either film badges or pocket dosimeters. A radiation intensity of 20 mr/hr was measured at the closest approach point of a man's torso to the installed liner, and an intensity of 60 mr/hr was measured at the closest approach point of a man's hands to the installed liner. The maximum average radiation dosage received by any operator during the experimental program was 60 mr/wk. This is 20% of the occupational dose rate which the AEC considers permissible.

Wear Determination Techniques

Radioassay techniques were developed to determine cylinder wear qualitatively and quantitatively. These radioassay procedures were developed for analyzing wear products distributed in both the lubricating oil and exhaust gas stream. Throughout the experimental program lubricating oil and exhaust gas samples were simultaneously collected for analysis at 6- to 16-hr intervals.

The precision of the radioactive tracer method was determined by comparing tracer-measured gross wear following each engine test with actual wear as determined from measurements using a bore gage micrometer. This was determined by measuring four diameters of the cylinder at eight elevations before and after each test. These readings were averaged to determine the diametrical increase at each elevation and thereby the total volume of material.

Lubricating Oil Radioassay

Two methods were investigated for the radioassay of lubricating oil samples. The first method involved various techniques for removing and concentrating radioactive debris from lube oil samples. An advantage of this procedure was the inherently high efficiency of gamma-ray detectors for concentrated radioactive debris. Only limited success was achieved in developing a suitable technique for concentrating wear product debris because of an inability to perform uniform separation of a high fraction of the radioactive debris from lube oil samples.

The second, and selected, method was the direct analysis of untreated lube oil samples. Special counting techniques conducive to extreme instrument stability, and counting procedures resulting in a maximum statistical counting efficiency were employed for direct lube oil sample analysis. The analysis of liter-volume samples with an immersion scintillation crystal-type detector, coupled with a precision scalar were employed for lube oil analysis.

Exhaust Gas Radioassay

An exhaust gas sampler designed for the isokinetic sampling of 0.4% of the exhaust gas stream of the test engine was installed in the exhaust system (Fig. 8). Since the volume of exhaust filtered was small, the duration of operation of the system varied depending on the actual debris carried by the exhaust gases. Debris collected on fiber glass filter

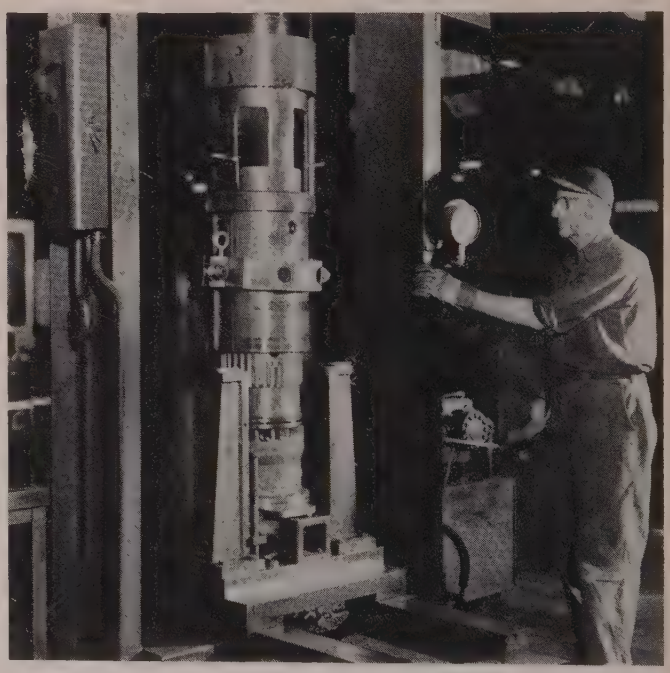


Fig. 5 — Liner jacket installation

mats was processed chemically for assaying in the well counter. It should be pointed out that this chemical separation procedure was necessitated by the capacity of existing equipment. To handle the large number of samples expected during the utilization of this technique, we are providing for a scintillation well counter with a large dimension crystal to permit direct counting of the exhaust filter mats.

Radioassay Standards

Cylinder-liner metal filings radioactiviated in both shielded and unshielded zones of the liner were processed for use as analytical standards. The properties of those standards were such that they would be subject to the same radioactive decay rates as was the cylinder-liner metal itself. These standards were radioassayed at periods corresponding to sample analysis, and used to compensate for the decay of the radioactive liner and for slight variations in instrument efficiency.

Data Processing

Knowing the engine lube oil consumption rate in addition to assay information, the actual wear can be calculated. These procedures were used in determining all data that are presented in this paper.

Experimental Program Results and Discussion

After the irradiated liner was installed in the engine assembly and the associated parts were secured, the engine was run-in to rated speed and load and a series of short runs were completed in order to check out the engine, instrumentation, and sampling equipment.

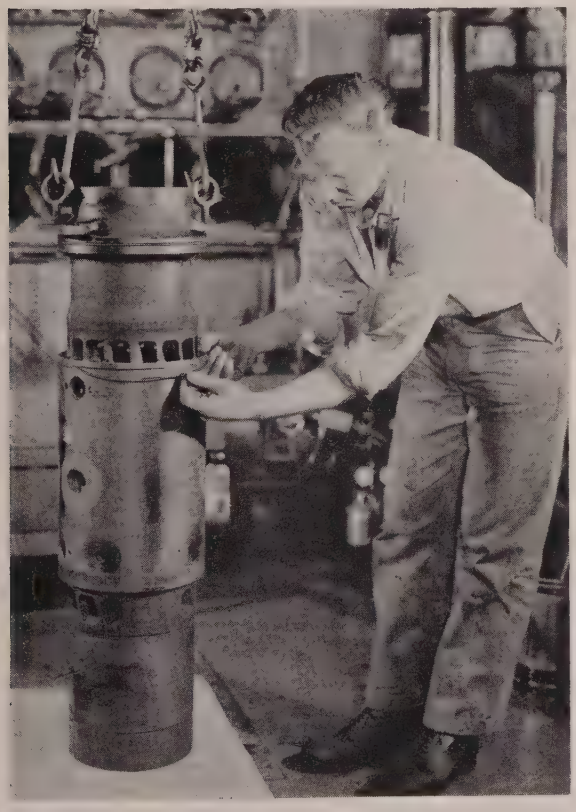


Fig. 6 — Engine buildup

Our purpose for the first radioactive cylinder liner was to determine the accuracy and reliability of the tracer technique for wear measurement. To do this, three tests were run for periods long enough to get good data with conventional bore gage instruments. At the end of each of these tests, the results with the bore gage were compared with those of the tracer technique. The first test was run on a No. 2 diesel fuel and resulted in a low rate of wear. The second test was run on a No. 6 residual fuel which had been compounded to give high rates of wear. No efforts were expended to reduce wear rates experienced during this test by use of premium lubricating oils or increased operating temperatures. The third test was actually a series of individual

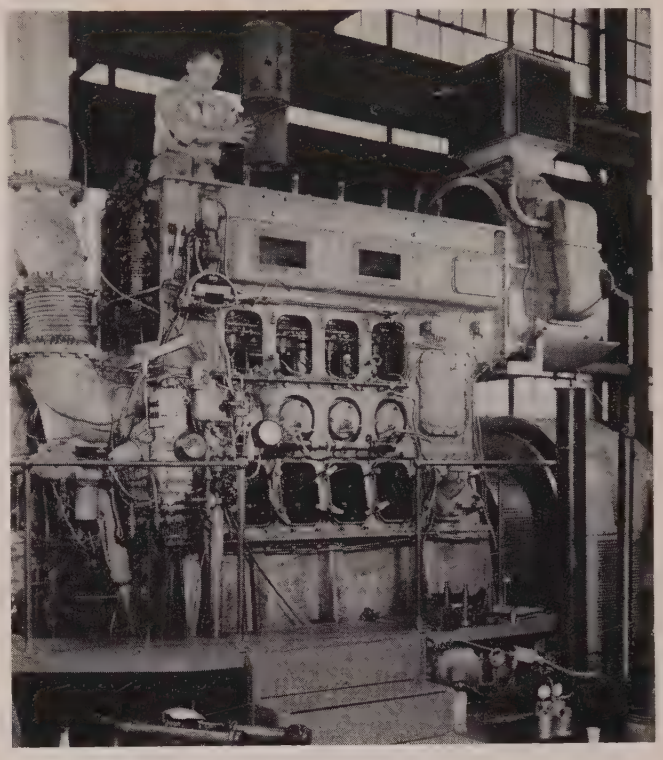


Fig. 7 — Engine buildup

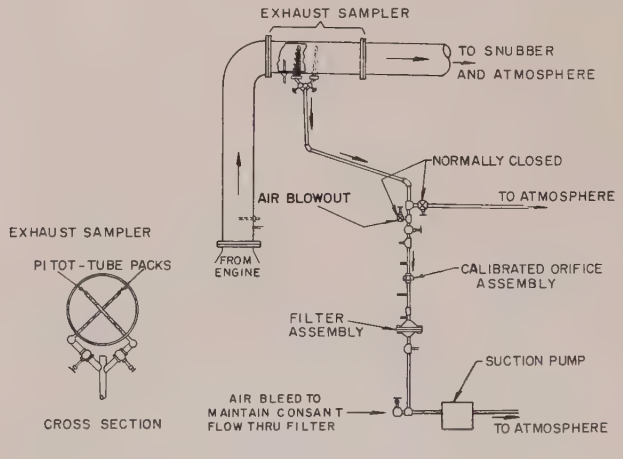


Fig. 8 — Exhaust gas sampling system

tests with a number of engine variables resulting in a variation of wear rates. It was our purpose in the third test to determine the ability of the tracer technique to detect relative wear rates quickly.

Discussion of Results

Figs 9–11 are plots of incremental wear data versus engine operating hours for Tests 1, 2, and 3. These curves show debris detected in the lube oil and in the exhaust gas stream plus deposit debris proportioned over the entire test period. This deposit debris, scraped from exhaust manifolds, liner combustion zone, oil strainer surfaces, and the rest, were required for calibration and will not be determined when utilizing this technique.

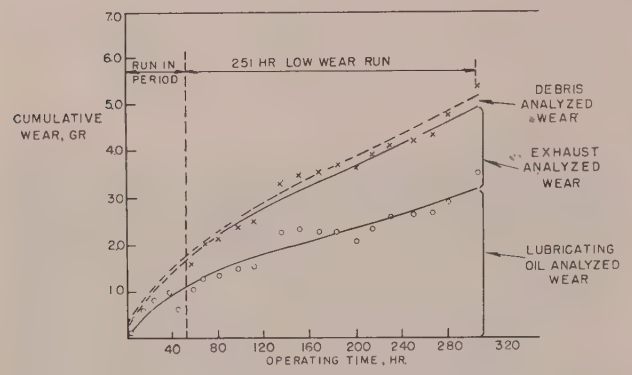


Fig. 9 — Cylinder-liner wear determined with radioactive tracer method for distillate fueled operation

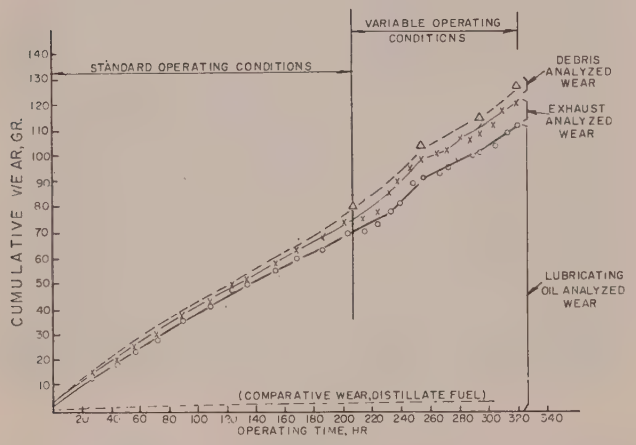


Fig. 10 — Cylinder-liner wear determined with radioactive tracer method for residual fueled operation

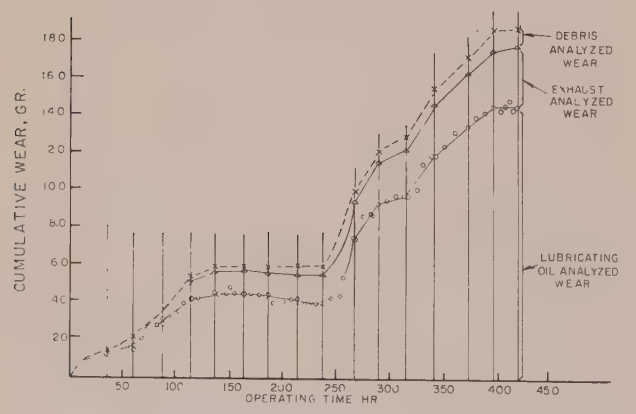


Fig. 11 — Cylinder-liner wear determined with radioactive tracers for variable conditions of engine operation

Conclusions

The important consideration of this investigation is not only what has been done, but what we are confident can be done in utilizing this technique in an effective manner. It seems quite reasonable that information never before obtainable can now be determined easily and relatively inexpensively. The effect on cylinder wear that is experienced with variables such as starting, idle, or cold temperature operation, sudden load changes, speed, torque, and the like, can be easily determined. Simultaneous data can be obtained on rings and cylinders, or other wearing parts of the engine through the use of different radioactive isotopes and pulse height analysis. It appears possible that local areas may be monitored through selective irradiation of the cylinder in conjunction with pulse height analysis.

Our confidence at Fairbanks, Morse and Co. in what this technique can achieve is such that we have obtained an AEC license, and have installed instrumentation to allow continuous monitoring of the lube oil and frequent checking of exhaust gas debris. We have acquired a second irradiated cylinder from the Battelle Memorial Institute for a continuing investigation of cylinder-liner wear. All parts and instruments are in process of installation, and initial testing will commence in the near future.

Fig. 12 compares the results of the three tests. The accuracy of the detection technique as well as the distribution of debris are clearly shown. It should be noted that the percentage of debris detected in the oil or exhaust gas varies considerably depending on various operating conditions. During Test No. 3, exhaust debris detection varied from 10 to 50% of the total.

The data from 305 hr of Test No. 1 and 210 hr of Test No. 2 were statistically analyzed for standard deviations. By using statistical methods commonly employed for analysis, it was possible to determine the reliability of the curves representing these data. The standard deviation of Test No. 1 was found to be 9.5% and Test No. 2, 5.4%.

Fig. 12 — Wear comparison for two methods of determination

Index—Vol. 68, 1960

AIRCRAFT ACCESSORIES AND EQUIPMENT
Ways of Improving Take-Off and Landing, F. W. Kolk 217
AIRCRAFT DESIGN
Preliminary Design Considerations for the Structure of a Trisonic Transport, M. G. Childers 396
Radiation Effects Studies for the Development of Radiation-Resistant Aircraft Subsystems, J. J. Tierney 634
AIRPORTS & AIRWAYS
The First Year of the Jet Age — Reflections, F. W. Kolk 578
Ways of Improving Take-Off and Landing, F. W. Kolk 217
ARNOLD, W. C.
Radioactive Cylinders — A Tool for Wear Research 699
(Co-authors V. T. Stonehocker, W. J. Braun, and D. N. Sunderman)
ARZT, P. R.
The Machining of Ultra-Strength Alloys 5
(Co-authors John Maranchik, Jr. and J. V. Gould)
ATOMIC ENERGY
High-Energy Accelerators, Keith R. Symon 157
BAUER, A. F.
Engine Blocks and Their Components in Aluminum Die Casting 385
BENNETT, P. A.
Reduction of Air Pollution by Control of Emission from Automotive Crankcases 514
(Co-authors M. W. Jackson, C. K. Murphy, and R. A. Randall)
BERCHTOLD, MAX
Road Performance of a Comprex Supercharged Diesel Truck .. 367
(Co-author H. P. Gull)
BIEHL, A. T.
Non-Air-Breathing Auxiliary Powerplants 111
(Co-authors D. L. Cochran, D. R. Sawle, M. R. Gustavson, and A. M. Taylor)
BLACKWOOD, A. K.
Correcting Horsepower Output — A Realistic Method for Diesel Engines 620
(Co-author W. J. McCulla)
BLOCH, PETER
Hydrodynamic Split Torque Transmissions 257
(Co-author Raymond C. Schneider)
BOWERS, E. H.
Hydrostatic Transmissions for Vehicles 248
BOZZELLI, A. J., JR.
Discussion of "Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance" by Theodore W. Selby 465
Discussion of "Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids" by H. E. Deen and C. M. Stendahl 465
BOZOS, C. P.
IH High-Speed Lightweight Diesels 501
(Co-authors A. F. Dewsberry and J. B. Reeves, Jr.)
BRACKEN, W. O.
Discussion of "The Expanding Polymer Horizon," by J. H. DuBois 84
BRAKES
Brake Drums — Investment for Safety and Profit, Paul G. Hykes and Clarence A. Herman 336
Economics of Heavy-Duty Brake Design and Maintenance, R. K. Super 346
Horsepower Rating for Commercial Vehicle Brakes, John H. King and Alex F. Stamm 646
BRAMER, SAUL E.
The Development of Refractory Sheet Metal Structures 352
(Co-author Alan V. Levy)
BRAUN, W. J.
Radioactive Cylinders — A Tool for Wear Research 699
(Co-authors W. C. Arnold, V. T. Stonehocker, and D. N. Sunderman)
BRUCKART, W. L.
Discussion of "The Development of Refractory Sheet Metal Structures" by Alan V. Levy and Saul E. Bramer 365
BRUHA, E. A.
Discussion of "Brake Drums — Investment for Safety and Profit" by Paul G. Hykes and Clarence A. Herman 349
Discussion of "Economics of Heavy-Duty Brake Design and Maintenance" by R. K. Super 349
BUSCARELLO, RALPH T.
Dynamic Balancing of Rebuilt Engines and Driveline Components 548
CHAMBERS, LESLIE A.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 529
CHIANG, C. W.
Physical and Chemical Ignition Delay in an Operating Diesel Engine Using the Hot-Motored Technique — Part II 562
(Co-authors P. S. Myers and O. A. Uyehara)
CHILDERS, M. G.
Preliminary Design Considerations for the Structure of a Trisonic Transport 396
CHRISTENSON, H. W.
Discussion of "Hydrostatic Transmissions" by P. C. Mortenson 256
Discussion of "Hydrostatic Transmissions for Vehicles" by E. H. Bowers 256
CHYLE, JOHN J.
Prediction in New Metal Joining Processes 143
COCHRAN, D. L.
Non-Air-Breathing Auxiliary Powerplants 111
(Co-authors A. T. Biehl, D. R. Sawle, M. R. Gustavson and A. M. Taylor)
COLEMAN, W. S.
Discussion of "Compressive Failures in Transmission Gearing" by G. E. Huffaker 58
COLYER, C. C.
Discussion of "Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance" by Theodore W. Selby 465
Discussion of "Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids" by H. E. Deen and C. M. Stendahl 465
COLEMAN, W. S.
Discussion of "Rating Transmissions from Highway Requirements and Vehicle Specifications" by F. C. Meldola 58
COMBUSTION, ABNORMAL — OTTO CYCLE ENGINE
Engine Rumble — A Barrier to Higher Compression Ratios?, R. F. Stebar, W. M. Wiese, and R. L. Everett 206
The Identification and Characteristization of Rumble and Thud, E. S. Starkman and W. E. Sytz 93
CONSTRUCTION & INDUSTRIAL EQUIPMENT
Radioisotopes in Research, W. P. Evans and E. W. Landen ... 19
Torsionally Resilient Drive Lines, P. J. Mazziotti 137
COOLEY, WILLIAM C.
Discussion of "Non-Air-Breathing Auxiliary Powerplants" by D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor 129
DALEY, J. A.
Radiation from Flames in Gas Turbine Combustors 554
(Co-authors R. M. Schirmer and L. A. McReynolds)
DAVISON, ELLARD D.
Developing Transaxle Fluid 685
(Co-authors N. A. Hunstad, R. A. Wilkins, and R. E. Osborne)
DEEN, H. E.
Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids 448
(Co-author C. M. Stendahl)
DEGA, ROBERT L.
Seal Testing To Establish Quality Control Specifications Can Reduce "Leakers" 440
(Co-author James D. Symons)
DEUBLE, NORMAN L.
Discussion of "The Development of Refractory Sheet Metal Structures" by Alan V. Levy and Saul E. Bramer 366
DEWSBERRY, A. F.
IH High-Speed Lightweight Diesels 501
(Co-authors C. P. Bozos and J. B. Reeves, Jr.)
DODGE, ADIEL Y.
Discussion of "Hydrodynamic Split Torque Transmissions" by Peter Bloch and Raymond C. Schneider 273
DU BOIS, J. H.
The Expanding Polymer Horizon 77
EATON, E. E.
Discussion of "Compressive Failures in Transmission Gearing" by G. E. Huffaker 58
Discussion of "Rating Transmissions from Highway Requirements and Vehicle Specifications" by F. C. Meldola 58
EGLI, HANS
Basic Design of Turbochargers for Diesel Engines 29
ELSE, H. D.
Discussion of "Non-Air-Breathing Auxiliary Powerplants" by D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor 128
ELECTRICAL EQUIPMENT (AIRCRAFT)
A Study of Self-Contained Starting Systems for Turbojet and Turboprop Engines, Henry R. Schmider and John H. Ferguson, Jr. 60
ENGINES, DIESEL
Basic Design of Turbochargers for Diesel Engines, Hans Egli 29
Correcting Horsepower Output — A Realistic Method for Diesel Engines, A. K. Blackwood and W. J. McCulla 620
Development of Two New Allis-Chalmers Diesel Engines, Hans L. Wittek 169
IH High-Speed Lightweight Diesels, A. F. Dewsberry, C. P. Bozos, and J. B. Reeves, Jr. 501
Physical and Chemical Ignition Delay in an Operating Diesel Engine Using the Hot-Motored Technique — Part II C. W. Chiang, P. S. Myers, and O. A. Uyehara 562
Radioactive Cylinders — A Tool for Wear Research, W. C. Arnold, V. T. Stonehocker, W. J. Braun, and D. N. Sunderman 699
Radiotracers Reveal Engine Wear during Detergent Oil Filtration Study, Harry Halliwell 288
Rolls-Royce Diesel Engines, R. H. Whiteside 101
A Universal Means for Rating Diesel Engines for Deposits and Wear, F. A. Robbins, L. G. Schneider, and G. H. Shea 653
The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear, B. A. Robbins, P. L. Pinotti, and D. R. Jones 298
ENGINES, GENERAL
GMR Stirling Thermal Engine, part of the Stirling Engine Story — 1960 Chapter, Gregory Flynn, Jr., Worth H. Percival, F. Earl Heffner 665
ENGINES, ROCKET
Electric Propulsion Systems, Victor P. Kovacik and Daniel P. Ross 380

Nuclear Rockets for Interplanetary Propulsion, Frank E. Rom and Paul G. Johnson 600
Some Development Problems with Large Cryogenic Propellant Systems, Daniel A. Heald 164
ENGINES, TURBOJET
Ways of Improving Take-Off and Landing, F. W. Kolk 217
EVANS, L. D.
Discussion of "Development of Two New Allis-Chalmers Diesel Engines" by Hans L. Wittek 188
EVANS, W. P.
Radioisotopes in Research 19
(Co-author E. W. Landen)
EVERETT, R. L.
Engine Rumble — A Barrier to Higher Compression Ratios? . 206
(Co-authors R. F. Stebar and W. M. Wiese)
EXHAUST SYSTEMS & EXHAUST GASES
Reduction of Air Pollution by Control of Emission from Automotive Crankcases, P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 514
FAITH, W. L.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 532
FARM MACHINERY
Soviet Agricultural Tractors and Equipment, Wayne H. Worthington 483
FERGUSON, JOHN H., JR.
A Study of Self-Contained Starting Systems for Turbojet and Turboprop Engines 60
(Co-author Henry R. Schmider)
FILIPPI, FRANK J.
Optimum Joint Design for High-Temperature Honeycomb Panels 658
(Co-author Boris Levenetz)
FISHER, D. P.
Discussion of "Horsepower Rating for Commercial Vehicle Brakes" by John H. King and Alex F. Stamm 651
FLECK, K. J.
Discussion of "Development of Two New Allis-Chalmers Diesel Engines" by Hans L. Wittek 188
FLEET OPERATION
Dynamic Balancing of Rebuilt Engines and Driveline Components, Ralph T. Buscarello 548
FLYNN, GREGORY, JR.
GMR Stirling Thermal Engine, part of the Stirling Engine Story — 1960 Chapter 665
(Co-authors Worth H. Percival and F. Earl Heifner)
FLYNN, H. O.
The Design of a Truck for Better Ride and Handling 408
FOX, ROBERT H.
Discussion of "Non-Air-Breathing Auxiliary Powerplants" by D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor 128
FORSTER, ING. HANS JOACHIM
Discussion of "Hydrodynamic Split Torque Transmissions" by Peter Bloch and Raymond C. Schneider 273
FRASS, A. P.
Discussion of "Non-Air-Breathing Auxiliary Powerplants" by D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor 128
FUELS, AVIATION
Radiation from Flames in Gas Turbine Combustors, R. M. Schirmer, L. A. McReynolds, and J. A. Daley 554
Turbine Fuel Thermal Stability, CFR Coker and Flight Evaluations, J. D. Rogers, Jr. 474
FUELS, DIESEL
The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear, B. A. Robbins, P. L. Pinotti, and D. R. Jones 298
FUELS, ROCKET
Current Trends in Storable Propellants, J. Silverman and S. A. Greene 282
GEARS
Computer-Based Selection of Balanced-Life Automotive Gears, G. J. Huebner 305
GOEPFRICH, R. A.
Discussion of "Brake Drums — Investment for Safety and Profit" by Paul G. Hykes and Clarence A. Herman 349
Discussion of "Economics of Heavy-Duty Brake Design and Maintenance" by R. K. Super 349
GOROWITZ, B.
Electric Propulsion Measurements with Small Thrust Plasma Generator 85
(Co-author B. W. Harned)
GOULD, J. V.
The Machining of Ultra-Strength Alloys 5
(Co-authors John Maranchik, Jr. and P. R. Arzt)
GREENE, D. F.
Discussion of "Developing Transaxle Fluid" by N. A. Hunstad, R. A. Wilkins, R. E. Osborne, and E. D. Davison 696
GREENE, S. A.
Current Trends in Storable Propellants 282
(Co-author J. Silverman)
GULL, H. P.
Road Performance of a Comprex Supercharged Diesel Truck . 367
(Co-author Max Berchtold)
GUSTAVSON, M. R.
Non-Air-Breathing Auxiliary Powerplants 111
(Co-authors D. L. Cochran, A. T. Biehl, D. R. Sawle, and A. M. Taylor)
HALLIWELL, HARRY
Radiotracers Reveal Engine Wear during Detergent Oil Filtration Study 288
HARKINS, J.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 531
HARNED, B. W.
Electric Propulsion Measurements with Small Thrust Plasma Generator 85
(Co-author B. Gorowitz)
HEALD, DANIEL A.
Some Development Problems with Large Cryogenic Propellant Systems 164
HEFFNER, F. EARL
GMR Stirling Thermal Engine, part of the Stirling Engine Story — 1960 Chapter 665
(Co-authors Gregory Flynn, Jr. and Worth H. Percival)
HEINEN, C. M.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 528
HERMAN, CLARENCE A.
Brake Drums — Investment for Safety and Profit 336
(Co-author Paul G. Hykes)
HOFMEISTER, W. F.
Application of the Cumulative Fatigue Damage Theory to Practical Problems 274
HUEBNER, G. J.
Computer-Based Selection of Balanced-Life Automotive Gears 305
HUFFAKER, G. E.
Compressive Failures in Transmission Gearing 53
HUNSTAD, NORMAN A.
Developing Transaxle Fluid 685
(Co-authors R. A. Wilkins, R. E. Osborne, and E. D. Davison)
HYKES, PAUL G.
Brake Drums — Investment for Safety and Profit 336
(Co-author Clarence A. Herman)
HYLER, JOHN
Discussion of "Electric Drive for Off-Highway Vehicles" by H. J. McLean and H. Vitt 241
ILKKA, GUST A.
Discussion of "New Elastomers Break Old Barriers" by Ralph P. Schmuckal 204
IGNITION SYSTEMS
Spark-Plug Misfiring — Mechanism Studies, H. P. Julien and R. F. Neblett 328
Transistor Switched Ignition Systems, G. E. Spaulding 610
JACKSON, H. R.
Discussion of "Radiotracers Reveal Engine Wear during Detergent Oil Filtration Study" by Harry Halliwell 296
JACKSON, M. W.
Reduction of Air Pollution by Control of Emission from Automotive Crankcases 514
(Co-authors P. A. Bennett, C. K. Murphy, and R. A. Randall)
JOHNSON, PAUL G.
Nuclear Rockets for Interplanetary Propulsion 600
(Co-author Frank E. Rom)
JONES, D. R.
The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear 298
(Co-authors B. A. Robbins and P. L. Pinotti)
JULIEN, H. P.
Spark-Plug Misfiring — Mechanism Studies 328
(Co-author R. F. Neblett)
KIBBEE, LEWIS C.
Discussion of "The Design of a Truck for Better Ride and Handling" by H. O. Flynn 438
KING, JOHN H.
Horsepower Rating for Commercial Vehicle Brakes 646
(Co-author Alex F. Stamm)
KIPP, K. L.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 531
KITE, WILLIAM H., JR.
Discussion of "Developing Transaxle Fluid" by N. A. Hunstad, R. A. Wilkins, R. E. Osborne, and E. D. Davison 697
KOENIG, E. F.
Discussion of "Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance" by Theodore W. Selby 466
Discussion of "Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids" by H. E. Deen and C. M. Stendahl 466
KOLK, F. W.
The First Year of the Jet Age — Reflections 578
Ways of Improving Take-Off and Landing 217
KOVACIK, VICTOR P.
Electric Propulsion Systems 380
(Co-author Daniel P. Ross)
KRAMER, A. O.
Discussion of "Road Performance of a Comprex Supercharged Diesel Truck" by Max Berchtold and H. P. Gull 378
KUIVINEN, T. O.
Discussion of "The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear" by B. A. Robbins, P. L. Pinotti, and D. R. Jones 303

LANDEN, E. W.
Radioisotopes in Research 19
(Co-author W. P. Evans)
LAWRASON, GEORGE C.
Discussion of "Radiotracers Reveal Engine Wear during Detergent Oil Filtration Study" by Harry Halliwell 296
LAZAN, B. J.
Material and Structural Damping for Vibration Control 537
LEE, WALTER J.
Possibility of "Spare Tire Elimination" by Use of Dual Compartment Tires 74
LETSINGER, J. H.
Discussion of "Compressive Failures in Transmission Gearing" by G. E. Huffaker 58
Discussion of "Rating Transmissions from Highway Requirements and Vehicle Specifications" by F. C. Meldola 58
LEVENETZ, BORIS
Optimum Joint Design for High-Temperature Honeycomb Panels 658
(Co-author Frank J. Filippi)
LEVY, ALAN V.
The Development of Refractory Sheet Metal Structures 352
(Co-author Saul E. Bramer)
LINSKY, BENJAMIN
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 535
LOEFFLER, B.
Discussion of "Development of Two New Allis-Chalmers Diesel Engines" by Hans L. Wittek 189
LUBRICANTS & LUBRICATION — CRANKCASE
Radiotracers Reveal Engine Wear during Detergent Oil Filtration Study, Harry Halliwell 288
LUBRICANTS & LUBRICATION — TRANSMISSION & AXLE
Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance, Theodore W. Selby 457
Developing Transaxle Fluid, N. A. Hunstad, R. A. Wilkins, R. E. Osborne, and E. D. Davison 685
Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids, H. E. Deen and C. M. Stendahl 448
MAGA, JOHN A.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 530
MALCOLMSON, R. W.
Discussion of "New Elastomers Break Old Barriers" by Ralph P. Schmuckal 205
MANLEY, L. W.
Discussion of "Developing Transaxle Fluid" by N. A. Hunstad, R. A. Wilkins, R. E. Osborne, and E. D. Davison 697
MARANCHIK, JOHN JR.
The Machining of Ultra-Strength Alloys 5
(Co-authors, J. V. Gould and P. R. Arzt)
MATERIALS, NONMETALLIC
Refractory Nonmetals for Use in Hypersonic Aircraft, John M. Nowak 571
MATTSON, R. L.
The Effect of Residual Stresses Induced by Strain-Peening upon Fatigue Strength 130
(Co-author J. G. Roberts)
MAZZIOTTI, P. J.
Torsionally Resilient Drive Lines 137
McCLINTOCK, R.
Discussion of "The Use of Radioactive Tracer Techniques to Determine the Effect of Operating Variables on Engine Wear" by B. A. Robbins, P. L. Pinotti, and D. R. Jones.... 303
McCORTNEY, W. J.
Discussion of "New Elastomers Break Old Barriers" by Ralph P. Schmuckal 205
McCULLA, W. J.
Correcting Horsepower Output — A Realistic Method for Diesel Engines 620
(Co-author A. K. Blackwood)
McGUIRE, R. C.
Discussion of "A Study of Self-Contained Starting Systems for Turbojet and Turboprop Engines" by Henry R. Schmider and John H. Ferguson, Jr. 73
McLEAN, H. J.
Electric Drive for Off-Highway Vehicles 232
(Co-author H. Vitt)
McMULLAN, O. W.
High-Temperature Materials and Their Heat-Treatment for Antifriction Applications 468
McREYNOLDS, L. A.
Radiation from Flames in Gas Turbine Combustors 554
(Co-authors R. M. Schirmer and J. A. Daley)
MELDOLA, F. C.
Rating Transmissions from Highway Requirements and Vehicle Specifications 49
MESSINA, N. V.
Discussion of "Developing Transaxle Fluid" by N. A. Hunstad, R. A. Wilkins, R. E. Osborne, and E. D. Davison 697
METALS, FERROUS
Application of the Cumulative Fatigue Damage Theory to Practical Problems, W. F. Hofmeister 274
METALS, NONFERROUS
The Development of Refractory Sheet Metal Structures, Alan V. Levy and Saul E. Bramer 352
Engine Blocks and Their Components in Aluminum Die Casting, A. F. Bauer 385
High-Temperature Materials and Their Heat-Treatment for Antifriction Applications, O. W. McMullan 468
MIAZGA, J. F.
Properties of Silicone Rubber for High-Temperature Static Seals 224
(Co-author H. E. Todd)
MISSILES
Non-Air-Breathing Auxiliary Powerplants, D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor .. 111
MORROW, JODEAN
Relaxation of Residual Stresses Due to Fatigue Loading.... 40
(Co-authors A. S. Ross and G. M. Sinclair)
MORTENSON, P. C.
Hydrostatic Transmissions 243
MURPHY, C. K.
Reduction of Air Pollution by Control of Emission from Automotive Crankcases 514
(Co-authors P. A. Bennett, M. W. Jackson, and R. A. Randall)
Discussion of "The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear" by B. A. Robbins, P. L. Pinotti, and D. R. Jones 303
MYERS, P. S.
Physical and Chemical Ignition Delay in an Operating Diesel Engine Using the Hot-Motored Technique — Part II 562
(Co-authors C. W. Chiang and O. A. Uyehara)
NEBLETT, R. F.
Spark-Plug Misfiring — Mechanism Studies 328
(Co-author H. P. Julien)
NOWAK, JOHN M.
Refractory Nonmetals for Use in Hypersonic Aircraft 571
NUCLEAR POWER
Non-Air-Breathing Auxiliary Powerplants, D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor.. 111
Nuclear Rockets for Interplanetary Propulsion, Frank E. Rom and Paul G. Johnson 600
Radiation Effects Studies for the Development of Radiation-Resistant Aircraft Subsystems, J. J. Tierney 634
OGDEN, E. B.
Discussion of "The Design of a Truck for Better Ride and Handling" by H. O. Flynn 438
OSBORNE, ROBERT E.
Developing Transaxle Fluid 685
(Co-authors N. A. Hunstad, R. A. Wilkins, and E. D. Davison)
PERCIVAL, WORTH H.
GMR Stirling Thermal Engine, part of the Stirling Engine Story — 1960 Chapter 665
(Co-authors Gregory Flynn, Jr. and F. Earl Heffner)
PERKINS, CHARLES M.
Discussion of "Compressive Failures in Transmission Gearing" by G. E. Huffaker 57
Discussion of "Rating Transmissions from Highway Requirements and Vehicle Specifications" by F. C. Meldola 57
PINOTTI, P. L.
The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear 298
(Co-authors B. A. Robbins and D. R. Jones)
PLASTICS & SYNTHETIC FIBERS
The Expanding Polymer Horizon, J. H. Du Bois 77
Properties of Asbestos Reinforced Laminates at Elevated Temperatures, Norman E. Wahl 627
POCOCK, ROBERT J.
Discussion of "Radiotracers Reveal Engine Wear during Detergent Oil Filtration Study" by Harry Halliwell 295
PRIZER, C. J.
Discussion of "Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance" by Theodore W. Selby 465
Discussion of "Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids" by H. E. Deen and C. M. Stendahl 465
PRODUCTION, METHODS
Prediction in New Metal Joining Processes, John J. Chyle..... 143
The Machining of Ultra-Strength Alloys, John Maranchik, Jr., J. V. Gould, and P. R. Arzt 5
Optimum Joint Design for High-Temperature Honeycomb Panels, Frank J. Filippi and Boris Levenetz 658
RAMSEL, C. A.
Discussion of "Hydrodynamic Split Torque Transmissions" by Peter Bloch and Raymond C. Schneider 273
RANDALL, R. A.
Reduction of Air Pollution by Control of Emission from Automotive Crankcases 514
(Co-authors P. A. Bennett, M. W. Jackson, and C. K. Murphy)
RASMUSSEN, WARREN
Discussion of "Seal Testing to Establish Quality Control Specifications Can Reduce 'Leakers'" by Robert L. Dega and James D. Symons 447
REEVES, J. B. JR.
IH High-Speed Lightweight Diesels 501
(Co-authors A. F. Dewsberry and C. P. Bozos)
ROBBINS, B. A.
The Use of Radioactive Tracer Techniques To Determine the Effect of Operating Variables on Engine Wear 298
(Co-authors P. L. Pinotti and D. R. Jones)
ROBBINS, F. A.
A Universal Means for Rating Diesel Engines for Deposits and Wear 653
(Co-authors L. G. Schneider and G. H. Shea)

ROBERTS, J. G.
The Effect of Residual Stresses Induced by Strain-Peening upon Fatigue Strength 130
(Co-author R. L. Mattson)
ROGERS, J. D. JR.
Turbine Fuel Thermal Stability, CFR Coker and Flight Evaluations 474
ROM, FRANK E.
Nuclear Rockets for Interplanetary Propulsion 600
(Co-author Paul G. Johnson)
ROSS, A. S.
Relaxation of Residual Stresses Due to Fatigue Loading 40
(Co-authors JoDean Morrow and G. M. Sinclair)
ROSS, DANIEL P.
Electric Propulsion Systems 380
(Co-author Victor P. Kovacik)
RUBIN, SHELDON
Design Procedure for Vibration Isolation on Nonrigid Supporting Structures 318
RUBBER & ELASTOMERS
New Elastomers Break Old Barriers, Ralph P. Schmuckal 193
Properties of Silicone Rubber for High-Temperature Static Seals, H. E. Todd and J. F. Miazga 224
SAWLE, D. R.
Non-Air-Breathing Auxiliary Powerplants 111
(Co-authors D. L. Cochran, A. T. Biehl, M. R. Gustavson, and A. M. Taylor)
SCHIRMER, R. M.
Radiation from Flames in Gas Turbine Combustors 554
(Co-authors L. A. McReynolds and J. A. Daley)
SCHMIDER, HENRY R.
A Study of Self-Contained Starting Systems for Turbojet and Turboprop Engines 60
(Co-author John H. Ferguson, Jr.)
SCHMUCKAL, RALPH P.
New Elastomers Break Old Barriers 193
SCHNEIDER, L. G.
A Universal Means for Rating Diesel Engines for Deposits and Wear 653
(Co-authors F. A. Robbins and G. H. Shea)
SCHNEIDER, RAYMOND C.
Hydrodynamic Split Torque Transmissions 257
SEALS (Co-author Peter Bloch)
Seal Testing To Establish Quality Control Specifications Can Reduce "Leakers," Robert L. Dega and James D. Symons.. 440
Properties of Silicone Rubber for High-Temperature Static Seals, H. E. Todd and J. F. Miazga 224
SELBY, THEODORE W.
Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance 457
SHEA, G. H.
A Universal Means for Rating Diesel Engines for Deposits And Wear 653
(Co-authors F. A. Robbins and L. G. Schneider)
SILVERMAN, J.
Current Trends in Storable Propellants 282
(Co-author S. A. Greene)
SINCLAIR, G. M.
Relaxation of Residual Stresses Due to Fatigue Loading 40
(Co-authors JoDean Morrow and A. S. Ross)
SPACE VEHICLES & SATELLITES
Electric Propulsion Measurements with Small Thrust Plasma Generator, B. Gorowitz and B. W. Harned 85
Non-Air-Breathing Auxiliary Powerplants, D. L. Cochran, A. T. Biehl, D. R. Sawle, M. R. Gustavson, and A. M. Taylor 111
SPAULDING, G. E. JR.
Transistor Switched Ignition Systems 610
SPINDT, R. S.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 534
STAMM, ALEX F.
Horsepower Rating for Commercial Vehicle Brakes 646
(Co-author John H. King)
STARKMAN, E. S.
The Identification and Characterization of Rumble and Thud.. 93
(Co-author W. E. Sytz)
STEBAR, R. F.
Engine Rumble — A Barrier to Higher Compression Ratios?.. 206
(Co-authors W. M. Wiese and R. L. Everett)
STENDAHL, C. M.
Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids 448
(Co-author H. E. Deen)
STONE, R. K.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 531
STONEHOCKER, V. T.
Radioactive Cylinders — A Tool for Wear Research 699
(Co-authors W. C. Arnold, W. J. Braun, and D. N. Sunderman)
STRESSES
Application of the Cumulative Fatigue Damage Theory to Practical Problems, W. F. Hofmeister 274
Relaxation of Residual Stresses Due to Fatigue Loading JoDean Morrow, A. S. Ross, and G. M. Sinclair 40
The Effect of Residual Stresses Induced by Strain-Peening upon Fatigue Strength, R. L. Mattson and J. G. Roberts 130
Optimum Joint Design for High-Temperature Honeycomb Panels, Frank J. Filippi and Boris Levenetz 658
SUNDERMAN, D. N.
Radioactive Cylinders — A Tool for Wear Research 699
(Co-authors W. C. Arnold, V. T. Stonehocker, and W. J. Braun)
SUPER, R. K.
Economics of Heavy-Duty Brake Design and Maintenance ... 346
SUPERCHARGING
Basic Design of Turbochargers for Diesel Engines, Hans Egli.. 29
Road Performance of a Comprex Supercharged Diesel Truck, Max Berchtold and H. P. Gull 367
SYMONS, JAMES D.
Seal Testing To Establish Quality Control Specifications Can Reduce "Leakers" 440
(Co-author Robert L. Dega)
SYMON, KEITH R.
High-Energy Accelerators 157
SYTZ, W. E.
The Identification and Characterization of Rumble and Thud.. 93
(Co-author E. S. Starkman)
TAYLOR, A. M.
Non-Air-Breathing Auxiliary Powerplants 111
(Co-authors D. L. Cochran, A. T. Biehl, D. R. Sawle, and M. R. Gustavson)
TEMPLIN, R. J.
Discussion of "Reduction of Air Pollution by Control of Emission from Automotive Crankcases" by P. A. Bennett, M. W. Jackson, C. K. Murphy, and R. A. Randall 532
TIERNEY, J. J.
Radiation Effects Studies for the Development of Radiation-Resistant Aircraft Subsystems 634
TILDEN, S. G.
Discussion of "Brake Drums — Investment for Safety and Profit" by Paul G. Hykes and Clarence A. Herman 350
Discussion of "Economics of Heavy-Duty Brake Design and Maintenance" by R. K. Super 350
Discussion of "Horsepower Rating for Commercial Vehicle Brakes" by John H. King and Alex F. Stamm 652
TINKLEPAUGH, J. R.
Discussion of "Refractory Nonmetals for Use in Hypersonic Aircraft" by John M. Nowak 577
TIRES, PASSENGER CAR
Possibility of "Spare Tire Elimination" by Use of Dual Compartment Tires, Walter J. Lee 74
TODD, H. E.
Properties of Silicone Rubber for High-Temperature Static Seals 224
(Co-author J. F. Miazga)
TRANSMISSIONS, HEAVY DUTY & TRUCK
Hydrodynamic Split Torque Transmissions, Peter Bloch and Raymond C. Schneider 257
Electric Drive for Off-Highway Vehicles, H. J. McLean and H. Vitt 232
Hydrostatic Transmissions, P. C. Mortenson 243
Hydrostatic Transmissions for Vehicles, E. H. Bowers 248
Rating Transmissions from Highway Requirements and Vehicle Specifications, F. C. Meldola 49
TRANSMISSIONS, PASSENGER CAR
Compressive Failures in Transmission Gearing, G. E. Huffaker 53
TRUCK DESIGN & PERFORMANCE
The Design of a Truck for Better Ride and Handling, H. O. Flynn 408
UITTI, PAUL M.
Discussion of "Road Performance of a Comprex Supercharged Diesel Truck" by Max Berchtold and H. P. Gull 379
UYEHARA, O. A.
Physical and Chemical Ignition Delay in an Operating Diesel Engine Using the Hot-Motored Technique — Part II 562
(Co-authors C. W. Chiang and P. S. Myers)
VANDERMAR, BERT
Discussion of "Automatic Transmission Fluid Viscosity at Low Temperature and Its Effect on Transmission Performance" by Theodore W. Selby 467
Discussion of "Oxidation Stability, Shear Stability, Rubber Swell Properties of Automatic Transmission Fluids" by H. E. Deen and C. M. Stendahl 467
VIBRATION, AIRCRAFT
Design Procedure for Vibration Isolation on Nonrigid Supporting Structures, Sheldon Rubin 318
Material and Structural Damping for Vibration Control, B. J. Lazan 537
VITT, H.
Electric Drive for Off-Highway Vehicles 232
(Co-author H. J. McLean)
WAHL, NORMAN E.
Properties of Asbestos Reinforced Laminates at Elevated Temperatures 627
WALTHER, W. D.
Discussion of "Transistor Switched Ignition Systems" by G. E. Spaulding, Jr. 619
WHITESIDE, R. H.
Rolls-Royce Diesel Engines 101
WIESE, W. M.
Engine Rumble — A Barrier to Higher Compression Ratios?.. 206
(Co-authors R. F. Stebar and R. L. Everett)
WILKINS, ROBERT A.
Developing Transaxle Fluid 685
(Co-authors N. A. Hunstad, R. E. Osborne, and E. D. Davison)
WITTEK, HANS L.
Development of Two New Allis-Chalmers Diesel Engines 169
WORTHINGTON, WAYNE H.
Soviet Agricultural Tractors and Equipment 483

DATE DUE

GAYLORD			PRINTED IN U.S.A.